Finite Random Matrix Theory Analysis of Multiple Antenna Communication Systems

POLITECNICO DI TORINO

SCUOLA DI DOTTORATO
Dottorato in Ingegneria Elettronica e delle Communicazioni
XXVII ciclo

Tesi di Dottorato

Finite Random Matrix Theory Analysis of Multiple Antenna Communication Systems

Siyuan Zhou
Mtr. 189424

Tutore
Prof. Carla-Fabiana Chiasserini

Coordinatore del corso di dottorato
Prof. Ivo Montrosset

March 2015

Summary

Multiple-antenna systems are capable of providing substantial improvement to wireless communication networks, in terms of data rate and reliability. Without utilizing extra spectrum or power resources, multiple-antenna technology has already been supported in several wireless communication standards, such as LTE, WiFi and WiMax. The surging popularity and enormous prospect of multiple-antenna technology require a better understanding to its fundamental performance over practical environments. Motivated by this, this thesis provides analytical characterizations of several seminal performance measures in advanced multiple-antenna systems. The analytical derivations are mainly based on finite dimension random matrix theory and a collection of novel random matrix theory results are derived.

The closed-form probability density function of the output of multiple-input multiple-output (MIMO) block-fading channels is studied. In contrast to the existing results, the proposed expressions are very general, applying for arbitrary number of antennas, arbitrary signal-to-noise ratio and multiple classical fading models. Results are presented assuming two input structures in the system: the independent identical distributed (i.i.d.) Gaussian input and a product form input. When the channel is fed by the i.i.d. Gaussian input, analysis is focused on the channel matrices whose Gramian is unitarily invariant. When the channel is fed by a product form input, analysis is conducted with respect to two capacity-achieving input structures that are dependent upon the relationship between the coherence length and the number of antennas. The mutual information of the systems can be computed numerically from the pdf expression of the output. The computation is relatively easy to handle, avoiding the need of the straight Monte-Carlo computation which is not feasible in large-dimensional networks.

The analytical characterization of the output pdf of a single-user MIMO block-fading channels with imperfect channel state information at the receiver is provided. The analysis is carried out under the assumption of a product structure for the input. The model can be thought of as a perturbation of the case where the statistics of the channel are perfectly known. Specifically, the average singular values of the channel are given, while the channel singular vectors are assumed to be isotropically distributed on the unitary groups of dimensions given by the number of transmit

and receive antennas. The channel estimate is affected by a Gaussian distributed error, which is modeled as a matrix with i.i.d. Gaussian entries of known covariance.

The ergodic capacity of an amplify-and-forward (AF) MIMO relay network over asymmetric channels is investigated. In particular, the source-relay and relay-destination channels undergo Rayleigh and Rician fading, respectively. Considering arbitrary-rank means for the relay-destination channel, the marginal distribution of an unordered eigenvalue of the cascaded AF channel is presented, thus the analytical expression of the ergodic capacity of the system is obtained. The results indicate the impact of the signal-to-noise ratio and of the Line-of-Sight component on such asymmetric relay network.

Acknowledgments

Foremost, I would like to express my deepest gratitude to my supervisor Prof. Carla-Fabiana Chiasserini, for her invaluable guidance and support in the past three years. Carla introduced me into the research field of multiple-antenna communication systems and provided me a constant flow of creative research topics. I am particularly thankful for her kindness in difficult times of the research, and for her permanent availability for my questions and discussions.

I would like to thank Alessandro Nordio and Giuseppa Alfano, who have helped me a lot during my whole PhD period. They provided me consistent support and helped me a lot especially in the initial period of research.

Special thanks goes to Carlo Borgiattino for his kindness and help, especially in the first year when I was a freshman in Politecnico. My thanks also goes to all members of TNG for their support and help in the past three years.

Finally, I will not forget the love and the support of my family through the time when I am abroad. Their love and encouragement empower me everyday.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Multiple-antenna communication technology is, by any measure, the most appealing topic and fastest evolving aspect in modern wireless communication systems. This technology is characterized by the employment of multiple antennas at either transmitter or receiver, or at both terminals. Compared with the conventional single-antenna system, multiple-antenna system is able to bring extra degree of freedom which can be applied to obtain various advantages including higher system capacity and more robust system performance.

Higher system capacity can be achieved by exploiting the spatial selectivity, without consuming extra power or bandwidth. The landmark papers [3] and [4] unveil such benefits of the multi-antenna system, which render a multiplication of the capacity. Actually, the original high-rate data sequence is split by the transmitter into multiple lower-rate data sequences, and then transmitted in parallel through each of the transmitter antennas, while the receiver undo the mixing of the MIMO channel in order to detect the signals corresponding to each of the transmitted data streams. This advantage is widely known as spatial multiplexing and it exists only in MIMO scenario.

On the other hand, in order to combat the fading and assure the reliability of the transmission, multiple-antenna system can also be exploited in terms of spatial diversity. By sending the same signals through multiple paths simultaneously, multiple independently faded replicas of data sequences are received and thus more reliable reception can be achieved. Differently to the temporal diversity or frequency diversity, which can be exploited in single-input single-output (SISO) system, the realization of spatial diversity does not incur a penalty in terms of data rate.

Of key importance for study of multiple-antenna systems is the characterization of the multipath transmission system. Modeling wireless channel is inherently difficult because the physical mechanisms of the channel are intricate to depict. It is also complicated to analyze a system with an involved channel model. Therefore, a few simplified canonical settings are put forth, which provide a compromise between

the realistic scenario and the tractability of the system model. If time is treated as the signaling domain, the description of the channel fading dynamics turns out to be a problem of the fading selectivity over each coded data packet. If the fading varies fast and ergodically during the packet length, i.e., the channel coherence length is much shorter than the packet length, the scenario can be regarded as the ergodic setting; if the fading is constant over the duration of one packet and varies only from packet to another, i.e., the channel coherence length is much longer than the packet length, the scenario can be treated as the quasi-static setting.

In ergodic setting, where one packet is subject to the entire distribution of the fading distribution, ergodic capacity is used to characterize the system performance. Ergodic capacity can be derived by taking expectation of the system mutual information over the stationary distribution of the fading. The error probability, which indicates the data rate is lower than the ergodic capacity, decays exponentially with the transmission length [5]. Ergodic capacity can be achieved by transmitting a codeword over a very large number of independent fading blocks. The most associated realistic scenario of the ergodic setting is the vehicular scenarios, where the mobile terminal moves at a high speed.

In quasi-static setting, each coded packet is subject to single realization of the fading process. Thus, the model is characterized as a block-fading structure, where the channel keeps constant during the block length. This situation typically occurs when stringent delay constraints are imposed, for example, in speech transmission over wireless networks, or when the channel varies slowly, for example, in indoor environments. Since this assumption gives rise to the nonexistence of the ergodic capacity, the relevant performance metric is the outage probability that characterizes the probability of a given bit rate that is not supported in a given fading channel. Furthermore, the concept of tradeoff between bit rate and outage probability has been raised [6] and it spurs a proliferation of research activities.

Currently, multiple-antenna technology is an essential element of multiple wireless communication standards such as IEEE 802.11n, IEEE 802.11ac, HSPA+, LTE-A, WiMAX, and WLAN. MIMO technology has been proven to be a prominent feature of future wireless communication systems.

1.1 MIMO Signals and Channel Models

The canonical form of MIMO system model is composed of multiple transmitter antennas and multiple receive antennas, as shown in Fig. 1.1. Let N_t denote the number of antennas at transmitter and N_r denote the number of antennas at receiver. Mathematically, the complex baseband model is characterized by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1.1}$$

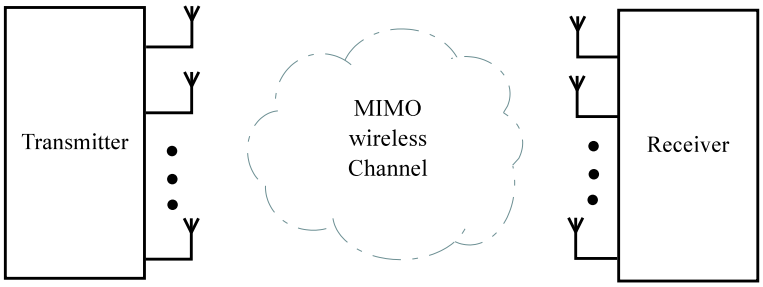

Figure 1.1: Diagram of a MIMO system

where

- $\mathbf{x}$ is an N_t dimension vector containing the complex signal transmitted by N_t antennas.
- $\mathbf{y}$ is an N_r dimension vector containing the complex signal received by N_r antennas.
- $\mathbf{n}$ is an N_t dimension vector representing the additive noise. Usually and also in this thesis, the elements of $\mathbf{n}$ are modeled as independent complex Gaussian random variables with zero mean and unit variance.
- $\mathbf{H}$ is an $N_r \times N_t$ dimension matrix, whose $(i,j)^{th}$ entry represents the multiplicative fading parameter for the wireless channel between the j^{th} transmit and i^{th} receive antenna.

Note that (1.1) applies directly to narrowband systems where the channel is subject to frequency flat fading. In narrowband system, the signal bandwidth is smaller than the coherence bandwidth of the channel. For more general wideband system, which appears more often in modern wireless communication systems, the channel is subject to frequency selective fading. By employing multi-carrier modulation scheme such as OFDM, the frequency selective channel can be decomposed into several parallel frequency flat channels. Therefore, the above model is still valid and worthy to be exploited even in wideband system.

The characteristic of the MIMO channel is determined by the statistical distribution of the entries of H. In realistic transmission, the transmitted signal is affected by both large-scale propagation effects and small-scale propagation effects. Large-scale propagation effects are characterized by path loss and shadowing, due to large transmission distances as well as large obstructing objects. Small-scale propagation effects are caused by the combination of constructive signals and destructive stochastic signals over multipath channels. These stochastic signals might

be scattered, reflected, diffracted, or delayed due to the variations of the channel. Particularly, small-scale fading is described as a stochastic process that undergoes Rayleigh distribution, Rician distribution, or other distributions. Throughout this thesis, we only consider small-scale propagation effects in the MIMO model. In the following, we first present mathematical description for three types of typical channel models considering the small-scale propagation, then we give a description of the block-fading channel.

1.1.1 Rayleigh Fading

In small-scale propagation model, Rayleigh fading is probably the most frequently used fading models. It is based on the assumption that a large number of propagation paths exist between the two ends, and the physical condition was referred by the name of rich scattering. The rich scattering assumption indicates the channel gains could be modeled as Rayleigh distribution. Mathematically, for any two Gaussian random variables X and Y, both of which are assumed zero mean and equal variance, it can be shown that $Z = \sqrt{X^2 + Y^2}$ is Rayleigh distributed. Rayleigh fading is proved to be a good match for multipath fading channels with no Line-of-Sight (LoS) path [4]. The channel fading amplitude z is distributed as

$$p(z) = \frac{z}{\sigma^2}\exp(-\frac{z^2}{2\sigma^2}), z \geq 0, \tag{1.2}$$

where $2\sigma^2$ is the average received signal power of the signal.

1.1.2 Rician Fading

If the channel has a fixed LoS component then the received signal equals the superposition of a complex Gaussian component and a LoS component. The channel fading amplitude z is distributed as [7]

$$p(z) = \frac{z}{\sigma^2}\exp\Big[\frac{-(z^2 + s^2)}{2\sigma^2}\Big] I_0(\frac{zs}{\sigma^2}), z \geq 0, \tag{1.3}$$

where $2\sigma^2$ is the average received power in the random multipath components, and s^2 is the power in the LoS component. The function I_0 is the modified Bessel function of 0th order. The Rician distribution is often described in terms of a fading parameter K, defined as [7]

$$K = \frac{s^2}{2\sigma^2}. \tag{1.4}$$

Therefore, K is the ratio of power in the determinant (LoS) component to the power in the stochastic (non-LoS) components. When $K = 0$, the Rician distribution turns out to be the Rayleigh distribution. When $K = \infty$, the channel is subject to no fading but only the LoS component [7].

1.1.3 Correlated Channel

Apart from the above two fading models, the transmission might be influenced by the spatial fading correlation. Spatial fading correlation may occur at either side of the channel, or both sides of the channel, mainly due to the insufficiently space between the antennas on the terminal. The correlated channel can be characterized through following *Kronecker* structure [17]

$$\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{H}_w\mathbf{S}^{\frac{1}{2}}, \tag{1.5}$$

where $\mathbf{R}$ and $\mathbf{S}$ are determinant correlation matrices, representing the correlation between the receive and transmit antennas, respectively.

1.1.4 Block-Fading Channel

The time-varying nature of the wireless channel is due to the movement of the transmitter or the receiver. In a fixed wireless environment where low Doppler spread occurs, the time between signal fades can be assumed to be sufficiently long. We call this scenario as slow fading where the coherence time is lower than the symbol period. The channel variation is slower than the baseband signal variation and the transmitter can send training signals that allow the receiver to estimate the channel accurately. In the other case, where the terminal is moving rapidly, for example in a high-speed train, the high Doppler spread occurs. This scenario is named as fast fading where coherence time is shorter than the symbol period.

In the scenario where at least one terminal is moving, the time between fades may be too short to permit reliable estimation of the channel. For instance, as stated in [8], A 60-mi/h mobile operating at 1.9 GHz has a fading interval of about 3 ms, which for a symbol rate of 30 kHz, corresponds to only about 100 symbol periods. In such 100 symbol periods, we could assume the channel coefficients are constant. Therefore, the block-fading channel is defined as the constant state of the wireless channel over a period of time that is called the coherence time (or coherence bandwidth if the system is analyzed in frequency domain). Particularly, for a rectangular Doppler spectrum, an exact relationship between the block-fading and the continuous-fading models is given in [9]

$$b = \frac{c}{2f_cT_sv} \tag{1.6}$$

where b, f_c and T_s stand for the coherence length, the carrier frequency and the symbol period, respectively. c and v represent the light speed and the velocity of the terminal. More elaborated description of the block-fading channel can be seen in (3.1).

1.2 Performance Measures

Based on the proposed MIMO models, several network performance measures are put forth in order to provide an accurate and simple means to evaluate the performance. These measures can also shed some light on the decision which has to be made when the engineers are designing or analyzing the networks, since they can easily uncover the impact of the system parameters on the network performance. In this section, we give a brief introduction of several essential measures which have been deeply exploited frequently in the literature.

1.2.1 Channel Capacity

Channel capacity is a measure which shows the maximum amount of information that can be transmitted and received with a negligible probability of error. The channel capacity is measured in bits/s/Hz. In MIMO system, due to the utilization of the space diversity, the data sequence can be split into several parallel independent sub-sequence. Hence MIMO system is able to provide a substantial improvement in terms of the channel capacity, compared with SISO system. From an information-theoretic point of view, what is of critical is the selectivity of the fading over the coded data block and two different notions of capacity emerged: the ergodic capacity and the outage capacity. If the fading channel varies greatly and ergodically during the transmission of the coded packet, ergodic capacity can be evaluated as the expectation of the mutual information between the input and the output over the distribution of the channel. For a system introduced in (1.1), the mutual information can be expressed as [3]

$$I = \log_2 \det\left(\mathbf{I} + \mathbf{H}\mathbf{\Phi}\mathbf{H}^\dagger\right) \tag{1.7}$$

where $\mathbf{\Phi}$ is the transmit covariance matrix defined by $\mathbf{\Phi} = E\{\mathbf{x}\mathbf{x}^\dagger\}$. Then the ergodic capacity can be obtained as [3]

$$C = E_{\mathbf{H}}\Big\{\max_{\mathrm{tr}\{\mathbf{\Phi}\}\leq P} I\Big\} \tag{1.8}$$

where P is the power constraint in the transmitter. If we consider that the transmitter has perfect CSI, the transmitter can allocate its power according to the well known waterfilling principle [10], which could maximize the mutual information. On the other hand, if the transmitter has no CSI, it is optimal to use a uniform power distribution [3]. We only consider the latter case. In this scenario, the transmit covariance matrix is then expressed as $\mathbf{\Phi} = \frac{P}{N_t}\mathbf{I}_{N_t}$, and the ergodic capacity is written as

$$C = E_{\mathbf{H}}\Big\{\max\Big(\log_2 \det\big(\mathbf{I} + \frac{\rho}{N_t}\mathbf{H}\mathbf{H}^\dagger\big)\Big)\Big\} \tag{1.9}$$

where ρ is the average SNR per receive antenna. Using eigenvalue decomposition, we can write $\mathbf{H}\mathbf{H}^{\dagger}$ as

$$\mathbf{H}\mathbf{H}^{\dagger} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^{\dagger} \tag{1.10}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues on the main diagonal, and $\mathbf{E}$ is the corresponding eigenvector matrix with orthogonal columns. Therefore, the ergodic capacity of MIMO system can be written as the sum of parallel AWGN SISO subchannels. The number of the subchannels is determined by the rank of $\mathbf{H}$, which is

$$\mathrm{rank}(\mathbf{H}) = r \leq \min\{N_t, N_r\}. \tag{1.11}$$

Replacing (1.10) into (1.9), the ergodic capacity can be expressed as

$$C = E_{\mathbf{H}}\Big\{ \sum_{i=1}^{r} \log_2\big(1 + \frac{\rho}{N_t}\lambda_i\big) \Big\} \tag{1.12}$$

However, the ergodic assumption does not always hold in realistic scenarios. For example, in indoor environment, there are no significant changes with respect to the fading channel over the span of each coded packet. In such case, it is possible that the transmission rate exceeds the instantaneous channel capacity. In order to depict the outage performance, $q\%$ outage capacity C_{out} indicates that, the mutual information I is guaranteed to be supported by $(100-p)\%$ of the channel realization, i.e.,

$$p(I \leq C_{\mathrm{out}}) = q\% \tag{1.13}$$

1.2.2 Signal-to-Noise Ratio and Signal-to-Interference-Plus-Noise Ratio

The SNR is usually measured at the receiver of the communication systems, where the co-channel interference is neglected. The output SNR depends explicitly on the channel and is defined as the ratio of the signal power to the noise power. In the multiuser scenario, if the co-channel interference is taken into account, the output SINR equals to the ratio of the signal power to the sum of the interference power and the noise power in the output. Apparently, the higher output SNR or SINR, the better the network performance. It should be noted that in the analysis of non-ergodic channel, the distribution of output SNR or SINR are essential measures to evaluate the system, since they can be used to express the outage probability or the average symbol error rate.

1.2.3 Outage Probability

Outage probability is another measure defined for the non-ergodic channel. It indicates the probability that the instantaneous channel capacity is below a specified

value, or the probability that the output SNR or SINR is lower than a threshold. Mathematically, outage probability can be obtained directly from the CDF of the SNR or SINR, as [7]

$$p_{\text{out}} = F_{\gamma}(\gamma_{\text{th}}) = \int_{0}^{\gamma_{\text{th}}} p_{\gamma}(\gamma)\mathrm{d}\gamma. \tag{1.14}$$

where $p_{\gamma}(\gamma)$ is the PDF of the SNR or SINR.

1.3 Related Work

Broadly speaking, compared with traditional SISO system, the most important benefits brought by MIMO systems can be categorized into two aspects: improved system capacity and enhanced link reliability. A large amount of work is based on the canonical system model introduced in (1.1). In this section, a concise description of such contributions is given. In addition to the classical model, more complicated and realistic network models are also proposed, which take into account the cooperative transmission or the block fading structure of the channel. A brief introduction is also provided regarding these models.

1.3.1 Capacity Based Research

The information-theoretic capacity based research on MIMO systems can be traced back to two landmark contributions in [3] and [4]. Both works uncover the fact that implementing multiple antennas on both sides of the channel can result in a multiplication of the capacity. This advantage is named as multiplexing gain. When the transmitter has no CSI, an effective transmission scheme has been presented in [11], called Vertical Bell-Labs Layered Space Time (VBLAST). The basic principle of such gain is to split a high-rate input sequence into multiple lower-rate sequences, which are sent through parallel channels by multiple transmit antennas, while the receiver undo and detect the mixing of the signal. For some systems that provide feedback link or a reciprocal channel for CSI, the transmitter is able to use the channel information to further improve the capacity. The optimal scheme for such application has been given in [12], named multichannel beamforming.

The majority of works on MIMO capacity select ergodic capacity as the measure of channel performance. The analysis is impacted heavily by the level CSI available on terminals. When the CSI is known perfectly both to the transmitter and to the receiver, the well known water-filling scheme is adopted to allocate the power among the transmitter antennas [3]. When the CSI is not available at transmitter and is perfectly known to the receiver due to the training process, even power distribution is the solution for the power allocation on the transmitter antenna. Although it is proved to be not strictly optimal [13], the justification for the even distribution as a

robust transmission scheme can be established based on the maxmin property [14]. Furthermore, power allocation is analyzed when the CSI is not perfectly know to the receiver and transmitter [15].

For certain specific MIMO channel models, such as uncorrelated Rayleigh fading scenario, the closed-form expression of ergodic capacity has been given in [16]. However, acquisition of the closed-form for capacity on more general fading models are much more involved. It is still an open problem to obtain the concise closed-form results for capacity on Rician or correlated models. The capacity of one-sided correlated Rayleigh fading channels has been studied in [17] and [18]. The capacity of Rician fading channels has been exploited in [19]. However, the obtained expressions of capacity in these works involve the summation of infinity number of entries, which are resulted from the expansion of the hypergeometric function. The main reason is that the eigenvalue distribution of $\mathbf{HH}^{\dagger}$ in (1.10) is more complicated for nonzero mean channel matrix or correlated channel matrix. In order to make progress in these cases as to provide engineering insights, some researches resort to bounding techniques or asymptotic results. In [20], tight upper and lower bounds for Rank-1 Rician channel has been derived. Regarding the correlated Rayleigh fading channel, both [21] and [16] have presented the upper bound and lower bound of the capacity, focusing on one-sided correlation and double-sided correlation, separately. The impact of correlation and LoS path have been considered simultaneously in [22], which provides the bound taking account of double-sided correlation and arbitrary rank channel means. Besides the bounding techniques, there is another approach to derive the asymptotic expression of ergodic capacity, assuming the number of transmit and receive antennas grows asymptotically while their ratio and SNR are constant. For example, [23] has adopted this approach and the asymptotic results are shown to be accurate even with a not-so-large dimension of the networks.

1.3.2 Diversity Based Research

Besides the capacity gain, another advantage brought by multiple-antenna systems is the spatial diversity gain. Diversity is an efficient approach to combat fading, and the general principle of diversity is transmitting multiple pieces of the same information through independent channels, so as to increase the robustness of the transmission, i.e. diversity gain. Mathematically, diversity gain is defined as the slope showing how fast the BER decays with the increase of the SNR. Diversity gain could be obtained by utilizing the temporal diversity or frequency diversity, by repeating the transmission in various time interval or the frequency interval. However, they will definitely impair the data rate or escalate the bandwidth usage. In contrast, spatial diversity will nor incur any penalty in data rate or bandwidth.

When CSI is not available at the transmitter, the same signal could be sent from single-antenna equipped transmitter through multiple independent channels.

The receiver can intelligently combine the multiple replicas of the same signal and a maximum diversity gain, which equals the number of receive antennas. When the transmitter is equipped with multiple antennas, a MIMO transmission scheme called space-time codes (STC) is proposed. Two classical STC techniques have been invented, which are space-time trellis codes [24] and space-time block codes [25, 26]. Both STC schemes are able to achieve the full diversity gain of MIMO system.

When CSI is available at the transmitter, a beamforming strategy called Maximum Ratio Transmission (MRT) is provided, which could fulfill the full diversity gain as well as the improved SNR performance [27]. The basic principle of MRT is transmit the signal along the eigenmode of the channel corresponding to the largest eigenvalue of the channel. After that, the receiver utilizes the Maximum Ratio Combining (MRC) scheme to combine the signal [28].

1.4 Dissertation Contributions and Outline

The rest of the thesis is organized as follows. Chapter 2 introduces mathematical background of the thesis, concentrating on the fundamental definition and theories of RMT. The following chapters present the major contributions of the thesis. In Chapter 3 and Chapter 4, we focus on the analysis of the output statistics of MIMO block-fading channels, while in Chapter 5, the ergodic capacity of amplify-and-forward relay MIMO network over asymmetric fading channels is analyzed.

Specifically, Chapter 3 considers the closed-form expression of the output pdf in MIMO block-fading channels. As the first work that focuses on the research of the output statistics of MIMO block-fading channels, two input structures and different channel fading models are considered in the network model. The analysis is based on a set of new statistics properties of the output corresponding to various system models using finite RMT. These novel output density expressions pave the way to compute the mutual information of the system, as well as other system performance measures such as the information density.

Chapter 4 exploits the output statistics of MIMO block-fading channels with imperfect CSI obtained at the receiver. The results are derived with the assumption that the estimated channel is impaired by the Gaussian noise. We derive the pdf of the received signal in several scenarios of practical interests, including the MMSE estimation. Furthermore, some open issues regarding the channel output characterization are presented.

Chapter 5 investigates the ergodic capacity performance of AF MIMO relay networks over asymmetric Rayleigh-Rician channels. In the two cases where the Rician channel has full-rank and low-rank means, the closed-form expression for the marginal pdf of an unordered eigenvalue of the cascaded AF channel is derived. Using these analytical expressions, the ergodic capacity of the system can be obtained.

Chapter 6 gives some concluding remarks and provides directions for future research.

Chapter 2

Mathematical Preliminaries

In this chapter, new results on the finite random matrix theory are presented. The new results pave the way for studying the output statistics of MIMO block-fading channels and the ergodic capacity of two-hop relay network over asymmetric channels in the subsequent chapters. In order to clarify the derivations in this thesis, a brief overview of finite RMT is first presented, which introduces the relevant concepts of determinant and generalized hypergeometric functions, as well as the matrix-variate distributions, joint eigenvalue and marginal eigenvalue distributions.

2.1 Basic Notations

We first enumerate the notations adopted in this thesis in Table.(2.1). Unless otherwise declared, throughout the thesis, uppercase and lowercase boldface letters denote matrices and vectors, respectively.

2.2 Definitions and Preliminary Results

This section provides the definitions, concepts and preliminary results, which are going to be used in subsequent sections and relevant derivations.

2.2.1 Determinant and Vandermonde Determinant

- *Determinant:* Let $\mathbf{A}$ be an $m \times m$ matrix with (i,j)th entry $a_{i,j}$, the determinant can be written as

$$|\mathbf{A}| = \sum_{\{\alpha\}} (-1)^{\mathrm{per}(\alpha)} \prod_{k=1}^{m} a_{k,p_k} \tag{2.1}$$

Table 2.1: Notations

$\sim$	Follows certain distribution.
$\in$	Belongs to.
$\sum$	Summation symbol.
$\prod$	Product symbol.
∞	Infinite symbol.
!	Factorial symbol.
$\rightarrow$	Approach symbol.
log, ln	Logarithm and natural logarithm.
e	Euler's constant.
max, min	Maximum and minimum.
lim	Limit.
$\mathbb{C}^m$	Complex $m \times 1$ vector.
$\mathbb{C}^{m\times n}$	Complex $m \times n$ matrix.
$\mathrm{tr}(\mathbf{X})$	Trace of the matrix $\mathbf{X}$.
$\mathrm{etr}(\mathbf{X})$	Shorthand for $\mathrm{e}^{\mathrm{tr}(\mathbf{X})}$.
$\lvert\mathbf{X}\rvert$,$\det(\mathbf{X})$	Determinant of matrix $\mathbf{X}$.
$\mathcal{V}(\mathbf{X})$	Vandermonde determinant of $\mathbf{X}$
$\{a_{i,j}\}$	The matrix whose elements are $a_{i,j}$
$\mathbf{I}$	Identity matrix.
$(\cdot)^T$,$(\cdot)^{\mathsf{H}}$	Transpose and conjugate transpose.
$\mathrm{diag}(a_1,\ldots,a_N)$	Square diagonal matrix with $a_1,\ldots,a_N$ on the diagonal.
$\lVert\cdot\rVert$	Euclidian norm, i.e.$\lVert\mathbf{X}\rVert^2 = \mathrm{Tr}(\mathbf{X}^{\mathsf{H}}\mathbf{X})$.
$\Gamma(\cdot)$	Standard gamma function.
$p(\cdot)$	Probability.
$K_v(\cdot)$	Modified Bessel function of the second kind.

where $p = \{p_1, \cdots, p_m\}$ is a permutation of $\{1, \cdots, m\}$ with sign $(-1)^{\text{per}(p)}$, and the sum is taken over all such permutations.

- *Determinant property:* Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times m$ and $m \times n$ matrices, respectively. Then
$$|\mathbf{I}_n + \mathbf{AB}| = |\mathbf{I}_n + \mathbf{BA}| \tag{2.2}$$

- *Determinant property:* Let both $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ matrices, then
$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}| = |\mathbf{BA}| \tag{2.3}$$

- *Complex multivariate Gamma function:* $\Gamma_m(a)$ is the complex multivariate Gamma function defined as [29]:
$$\Gamma_m(a) = \pi_m \prod_{\ell=1}^{m} \Gamma(a - \ell + 1) \tag{2.4}$$
where m being a non-negative integer and
$$\pi_m = \pi^{m(m-1)/2} \tag{2.5}$$

- *Vandermonde determinant:* Let $\mathbf{A}$ be an $m \times m$ Hermitian matrix with eigenvalues $a_1, \ldots, a_m$. Then the Vandermonde determinant of $\mathbf{A}$ is defined as [30, eq. (2.10)]:
$$\mathcal{V}(\mathbf{A}) = \prod_{1 \le i < j \le m} (a_i - a_j)\,, \tag{2.6}$$
where we assume the eigenvalues to be ordered in decreasing order so that $\mathcal{V}(\mathbf{A})$ is non negative. Moreover, for any constant c, we have $\mathcal{V}(c\mathbf{A}) = c^{m(m-1)/2}\mathcal{V}(\mathbf{A})$.

- *Vandermonde determinant property:* Let $\mathbf{F} = \{f_i(a_j)\}$, $i,j = 1, \ldots, m$, be an $m \times m$ matrix, where $f_i(\cdot)$'s are any differentiable functions. Clearly, if the eigenvalues of $\mathbf{A}$ are not distinct, $\mathcal{V}(\mathbf{A}) = 0$ and $|\mathbf{F}| = 0$. In such a case, the ratio $|\mathbf{F}|/\mathcal{V}(\mathbf{A})$, which appears in the density of many matrices that we study in the following, can be evaluated by applying l'Hôpital's rule. More precisely, let n be an integer such that $0 < n < m$, then [31, Lemma 5]
$$\lim_{a_{n+1}, \ldots, a_m \to a} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{A})} = \frac{\pi_m \Gamma_n(m)}{\pi_n \Gamma_m(m)} \frac{|\widetilde{\mathbf{F}}|}{\mathcal{V}(\widetilde{\mathbf{A}})} |\widetilde{\mathbf{A}} - a\mathbf{I}|^{n-m} \tag{2.7}$$
where $\widetilde{\mathbf{A}}$ is of size $n \times n$ and has eigenvalues $a_1, \ldots, a_n$ and
$$(\widetilde{\mathbf{F}})_{ij} = \begin{cases} f_i(a_j) & i = 1, \ldots, m;\ j = 1, \ldots, n \\ f_i^{(m-j)}(a) & i = 1, \ldots, m;\ j = n+1, \ldots, m \end{cases} \tag{2.8}$$

with $f_i^{(k)}(\cdot)$ denoting the k-th derivative of $f_i(\cdot)$. For $n = 0$, we have

$$\lim_{a_1,\ldots,a_m \to a} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{A})} = \frac{\pi_m}{\Gamma_m(m)} |\widetilde{\mathbf{F}}| \tag{2.9}$$

where $(\widetilde{\mathbf{F}})_{ij} = f_i^{(m-j)}(a)$, $i,j = 1,\ldots,m$.

2.2.2 Generalized Hypergeometric Function

The generalized hypergeometric function is defined as ${}_pF_q(\mathbf{a};\mathbf{b};\mathcal{X})$, where $\mathbf{a} = [a_1,\ldots,a_p]^\mathsf{T}$, $\mathbf{b} = [b_1,\ldots,b_q]^\mathsf{T}$, and $\mathcal{X}$ is a set of arguments that can be either scalars or square matrices [32]. For the arguments with matrix, the definition of zonal polynomial of a matrix argument is first provided, since it plays an important role in the original definition of hypergeometric function with matrix arguments.

- *Hypergeometric function with scalar argument:* In the case of a single scalar argument, $\mathcal{X} = \{x\}$, the generalized hypergeometric function is defined as in [30, eq. (2.24)]:

$$ {}_pF_q(\mathbf{a};\mathbf{b};x) = \sum_{k=0}^{\infty} \frac{[\mathbf{a}]_k}{[\mathbf{b}]_k} \frac{x^k}{k!} \tag{2.10}$$

  where $[\mathbf{a}]_k = \prod_{i=1}^{p} [a_i]_k$, $[\mathbf{b}]_k = \prod_{j=1}^{q} [b_j]_k$, and $[z]_k = \Gamma(z+k)/\Gamma(z)$ denotes the Pochhammer symbol. Note that ${}_0F_0(;;x) = \mathrm{e}^x$, and ${}_1F_0(a;;x) = (1-x)^{-a}$. The function ${}_0F_1(;b;x)$ is closely related to the Bessel's function, and in the literature functions ${}_1F_1(a;b;x)$ and ${}_2F_1(a_1,a_2;b;x)$ are also called *confluent hypergeometric function of the first kind* and *Gauss's hypergeometric function*, respectively.

- *Zonal polynomials:* Zonal polynomials are an essential component in the expression of hypergeometric function with matrix argument. Here we introduce the complex zonal polynomials of Hermitian matrix arguments. Assuming $\mathbf{X}$ is a $n \times n$ Hermitian matrix, the complex zonal polynomial $\tilde{C}_\mathcal{K}(\mathbf{X})$ is defined as [29]

$$\tilde{C}_\mathcal{K}(\mathbf{X}) = \chi_{[\mathcal{K}]}(1)\chi_{\{\mathcal{K}\}}(\mathbf{X}) \tag{2.11}$$

  where $\chi_{\{\mathcal{K}\}}(\mathbf{X})$ is the character of the representation $\{\mathcal{K}\}$ of the linear group, given as a symmetric function of the eigenvalues $x_1,\ldots,x_n$ of $\mathbf{X}$ by

$$\chi_{\{\mathcal{K}\}}(\mathbf{X}) = \frac{\left|\{x_i^{k_j+n-j}\}_{i,j=1,\ldots,n}\right|}{\left|\{x_i^{n-j}\}_{i,j=1,\ldots,n}\right|} \tag{2.12}$$

  and $\chi_{[\mathcal{K}]}(1)$ is the dimension of the representation $[\mathcal{K}]$ of the symmetric group given by

$$\chi_{[\mathcal{K}]}(1) = k! \frac{\prod_{i<j}^{n}(k_i - k_j - i + j)}{\Gamma_n(n,\mathcal{K})} \tag{2.13}$$

where

$$\Gamma_n(n,\mathcal{K}) = \prod_{i=1}^{n} \Gamma(n + k_i - i + 1) \tag{2.14}$$

- *Hypergeometric function with one matrix argument:* As defined in [33], assuming $\mathbf{X}$ is a Hermitian $n \times n$ matrix, the hypergeometric function of a matrix argument ${}_pF_q(\mathbf{a};\mathbf{b};\mathbf{X})$ is defined as

$${}_pF_q(\mathbf{a};\mathbf{b};\mathbf{X}) = \sum_{k=0}^{\infty} \sum_{\mathcal{K}} \frac{[\mathbf{a}]_k}{[\mathbf{b}]_k} \frac{\tilde{C}_{\mathcal{K}(\mathbf{X})}}{k!} \tag{2.15}$$

where $\mathcal{K} = (k_1,\ldots,k_n)$ is a partition of k, and $(\cdot)_{\mathcal{K}}$ is the complex multivariate hypergeometric coefficient

$$(a)_{\mathcal{K}} = \prod_{l=1}^{n} (a - l + 1)_{k_l} \tag{2.16}$$

- *Hypergeometric function with two matrix arguments:* As defined in [29], the hypergeometric function of two matrix arguments is defined as follows

$${}_pF_q(\mathbf{a};\mathbf{b};\mathbf{X},\mathbf{Y}) = \sum_{k=0}^{\infty} \sum_{\mathcal{K}} \frac{[\mathbf{a}]_k}{[\mathbf{b}]_k} \frac{\tilde{C}_{\mathcal{K}(\mathbf{X})} \tilde{C}_{\mathcal{K}(\mathbf{Y})}}{\tilde{C}_{\mathcal{K}(\mathbf{I})} k!} \tag{2.17}$$

The main problem of (2.15) and (2.17) is the extreme difficulty to numerical compute the hypergeometric function. This issue is resulted from the infinite summation of partitions. Therefore, the expansion with zonal polynomial could not provide a clear guidance to the system design. Another expansion approach, called determinant representation, is given with particularly compact and handy form compared with the zonal polynomial expansion. The generalized hypergeometric function of two matrix arguments, $\mathcal{X} = \{\mathbf{\Phi},\mathbf{\Psi}\}$, both of size $m \times m$, can be written through hypergeometric functions of scalar arguments as [30, eq. (2.34)]

$${}_pF_q(\mathbf{a};\mathbf{b};\mathbf{\Phi},\mathbf{\Psi}) = c \frac{|\{{}_pF_q(\tilde{\mathbf{a}};\tilde{\mathbf{b}};\phi_h\psi_k)\}|}{\mathcal{V}(\mathbf{\Phi})\mathcal{V}(\mathbf{\Psi})} \tag{2.18}$$

$h,k = 1,\ldots,m$, where the constant c is given by [30]

$$c = \frac{\Gamma_m(m)}{\pi_m^{q-p+1}} \left[\prod_{j=1}^{q} \frac{\Gamma_m(b_j)}{(b_j - m)!^m}\right] \left[\prod_{i=1}^{p} \frac{(a_i - m)!^m}{\Gamma_m(a_i)}\right] \tag{2.19}$$

$\tilde{a}_i = a_i - m + 1$, $i = 1,\ldots,p$, $\tilde{b}_j = b_j - m + 1$, $j = 1,\ldots,q$, and the eigenvalues of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are denoted by $\phi_1,\ldots,\phi_m$ and $\psi_1,\ldots,\psi_m$, respectively.

- *Derivative of hypergeometric function:* The ℓ-th derivative of the generalized hypergeometric function ${}_pF_q(\mathbf{a};\mathbf{b};sx)$ is given by [32]:

$$\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}{}_pF_q(\mathbf{a};\mathbf{b};sx) = s^\ell \frac{(\mathbf{a})_\ell}{(\mathbf{b})_\ell}{}_pF_q(\tilde{\mathbf{a}};\tilde{\mathbf{b}};sx) \tag{2.20}$$

  where $\tilde{a}_i = a_i + \ell$, $\tilde{b}_j = b_j + \ell$, $i = 1,\ldots,p$, $j = 1,\ldots,q$, and s is a parameter.

2.3 Matrix-Variate Distributions

In this section, we present multiple distribution expressions that we are going to utilize in this thesis; namely, matrix-variate complex Gaussian distribution, complex Wishart distribution and eigenvalue distribution. Before giving such expressions, we first present the definitions of Stiefel matrix, Haar matrix and the property of unitarily invariant. Note that the matrix that we deal with in this thesis are all referred to complex matrix. So from now on we will neglect the specification of complex entries of the matrix.

- *Stiefel matrix:* An $m \times n$ $(m \geq n)$ random *Stiefel* matrix $\mathbf{S} \in \mathcal{S}(m,n)$ is such that $\mathbf{S}^\mathsf{H}\mathbf{S} = \mathbf{I}$ and is uniformly distributed on $\mathcal{S}(m,n)$. Then, it has pdf $p(\mathbf{S}) = |\mathcal{S}(m,n)|^{-1}$.

- *Haar matrix:* A square $m \times m$ random Haar matrix (Unitary matrix) $\mathbf{U} \in \mathcal{U}(m)$ is such that $\mathbf{U}\mathbf{U}^\mathsf{H} = \mathbf{U}^\mathsf{H}\mathbf{U} = \mathbf{I}$. When it is uniformly distributed on $\mathcal{U}(m)$, it has pdf $p(\mathbf{U}) = |\mathcal{U}(m)|^{-1}$.

- *Unitarily invariant:* An $m \times m$ Hermitian random matrix $\mathbf{A}$ is unitarily invariant if the joint distribution of its entries equals that of $\mathbf{V}\mathbf{A}\mathbf{V}^\mathsf{H}$ where $\mathbf{V}$ is any unitary matrix independent of $\mathbf{A}$ [34, Definition 2.6 and Lemma 2.6]. If $\mathbf{A}$ is unitarily invariant, then its eigenvalue decomposition can be written as $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\mathsf{H}$ where $\mathbf{U}$ is a Haar matrix independent of the diagonal matrix $\mathbf{\Lambda}$. Since $\mathbf{U}$ is Haar (isotropic), it is uniformly distributed on $\mathcal{U}(m)$.

- *Beta-distribution:* The $n \times n$ random matrix $\mathbf{B}$ is Beta-distributed with positive integer parameters p and q ($\mathbf{B} \sim \mathcal{B}_n(p,q)$) if

  - given $\mathbf{T}$ an upper triangular matrix with positive diagonal elements, we can write $\mathbf{B} = (\mathbf{T}^\mathsf{H})^{-1}\mathbf{C}\mathbf{T}$ where $\mathbf{C} \sim \mathcal{W}_n(p,\mathbf{\Theta})$, and
  - given $\mathbf{A} \sim \mathcal{W}_n(m,\mathbf{\Theta})$, we can write $\mathbf{A} + \mathbf{C} = \mathbf{T}^\mathsf{H}\mathbf{T}$. Notice that, if either $p < n$ or $q < n$, or both $p < n$ and $q < n$, the distribution is referred to as *pseudo*-Beta since it involves *pseudo*-Wishart matrices [2, and references therein].

2.3.1 Matrix-Variate Complex Gaussian Distribution

Let $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{M},\mathbf{\Omega}\otimes\mathbf{\Sigma})$ denote the matrix-variate complex Gaussian distribution of $\mathbf{X}$. We have the pdf of $\mathbf{X}$ as [29]

$$p(\mathbf{X}) = \frac{\text{etr}\Big(-\mathbf{\Omega}^{-1}(\mathbf{X}-\mathbf{M})\mathbf{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})^{\mathsf{H}}\Big)}{\pi^{nm}|\mathbf{\Omega}|^{m}|\mathbf{\Sigma}|^{n}} \tag{2.21}$$

where $\mathbf{M}$ is the mean matrix, $\mathbf{\Omega}$ and $\mathbf{\Sigma}$ are covariance matrices where $\mathbf{\Omega} \in \mathcal{C}^{n\times n} > 0$ and $\mathbf{\Sigma} \in \mathcal{C}^{m\times m} > 0$.

2.3.2 Complex Wishart Distribution

- *Wishart matrix:* If the columns of the $m\times n$ matrix $\mathbf{H}$ are independent complex Gaussian vectors with covariance $\mathbf{\Theta}$, assuming $m \leq n$, the $m \times m$ random matrix $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$ is called Wishart matrix.

- *Central Wishart matrix distribution:* If matrix $\mathbf{H}$ is zero mean, the Wishart matrix $\mathbf{W}$ is named central Wishart matrix as $\mathbf{W} \sim \mathcal{W}_m(n,\mathbf{\Theta})$. For $m \leq n$, the PDF of $\mathbf{W}$ is given as [34]

$$p(\mathbf{W}) = \frac{\pi^{-m(m-1)}/2}{|\mathbf{\Theta}|^{n}\prod_{i=1}^{m}(n-i)!}\exp[-\text{tr}\{\mathbf{\Sigma}^{-1}\mathbf{W}\}]|\mathbf{W}|^{n-m} \tag{2.22}$$

- *Non-central Wishart matrix distribution:* If matrix $\mathbf{H}$ is with mean $\mathbf{M} = \mathbb{E}[\mathbf{H}]$, matrix $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$ is non-central Wishart [29]. For $m \leq n$, the distribution of $\mathbf{W}$ is given by [29]

$$p(\mathbf{W}) = \frac{|\mathbf{W}|^{n-m}}{\Gamma_m(n)}\frac{{}_0F_1(\,;n;\mathbf{M}\mathbf{M}^{\mathsf{H}}\mathbf{W})}{\mathrm{e}^{\mathsf{Tr}\{\mathbf{W}+\mathbf{M}\mathbf{M}^{\mathsf{H}}\}}}\,. \tag{2.23}$$

When $m \geq n$, the same expression as in (2.22) and (2.23) hold but replacing $\mathbf{M}$, m and n with, respectively, $\mathbf{M}^{\mathsf{H}}$, n and m.

2.3.3 Eigenvalue Distribution

Many results presented in this thesis are dependent on the eigenvalue distribution of random matrices. Among them, the eigenvalue distributions of Wishart matrix are fundamental to the derivation of system performance of MIMO systems. We will present the joint eigenvalue distribution of central Wishart matrix and non-central Wishart matrix.

- For $m \leq n$, let $\mathbf{W} = \mathbf{U}\mathit{\Lambda}\mathbf{U}^{\mathsf{H}}$ be the singular value decomposition (SVD) of $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$, where $\mathbf{H}$ is with zero mean. If $\mathbf{\Theta} = \mathbf{I}$, then $\mathbf{W}$ is unitarily invariant [34]. In such a case, the joint distribution of the ordered eigenvalues $\mathit{\Lambda}$ can be written as [35, 29]

$$p(\mathbf{\Lambda}) = \frac{\pi_m^2 |\mathbf{\Lambda}|^{n-m} \mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Lambda}\}}}{\Gamma_m(n)\Gamma_m(m)} \mathcal{V}^2(\mathbf{\Lambda}) \,. \tag{2.24}$$

- For $m > n$, if the rows of $\mathbf{H}$ are independent and their covariance matrix is $\mathbf{I}$, the distribution of the ordered eigenvalues of $\mathbf{H}^{\mathsf{H}}\mathbf{H}$ is given by [35]

$$p(\mathbf{\Lambda}) = \frac{\pi_{\mathbf{n}}^{\mathbf{2}} |\mathbf{\Lambda}|^{\mathbf{m}-\mathbf{n}} \mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Lambda}\}}}{\Gamma_{\mathbf{n}}(\mathbf{m})\Gamma_{\mathbf{n}}(\mathbf{n})} \mathcal{V}^{\mathbf{2}}(\mathbf{\Lambda}) \,. \tag{2.25}$$

- For $m \leq n$, let $\mathbf{W} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$ be the singular value decomposition (SVD) of $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$, where $\mathbf{H}$ is with zero mean. If $\mathbf{\Theta}$ has distinctive nonzero eigenvalues, then the joint distribution of the ordered eigenvalues $\mathbf{\Lambda}$ can be written as [29]

$$p(\mathbf{\Lambda}) = \frac{\pi_m^2 |\mathbf{\Lambda}|^{n-m} \mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Lambda}\}}}{\Gamma_m(m)} \mathcal{V}^2(\mathbf{\Lambda}) |\mathbf{\Theta}|^{-n} {}_0F_0(-\mathbf{\Theta}^{-1}, \mathbf{\Lambda}) \,. \tag{2.26}$$

- For $m \leq n$, let $\mathbf{H}$ be an $m \times n$ random matrix whose entries are independent, Gaussian random variables with unit variance and average $\mathbf{M} = \mathbb{E}[\mathbf{H}]$. If $\mathbf{M}\mathbf{M}^{\mathsf{H}}$ has full rank and distinct eigenvalues, $\mu_1, \ldots, \mu_m$, then the joint pdf of the ordered, strictly positive eigenvalues $(\lambda_1, \ldots, \lambda_m) = \mathrm{diag}(\mathbf{\Lambda})$ of $\mathbf{W}$ is given by [29, eq. (102)]

$$p(\mathbf{\Lambda}) = \frac{|\mathbf{\Lambda}|^{n-m} \mathcal{V}(\mathbf{\Lambda}) |\{ {}_0F_1(\,;n-m+1;\mu_i\lambda_j)\}|}{(n-m)!^m \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Lambda}+\mathbf{M}\mathbf{M}^{\mathsf{H}}\}} \mathcal{V}(\mathbf{M}\mathbf{M}^{\mathsf{H}})} \,. \tag{2.27}$$

  Note that (2.27) has been obtained from [29, eq. (102)] by exploiting the result in (2.18).

  As can be observed from (2.23), if $\mathbf{M}\mathbf{M}^{\mathsf{H}}$ is a scalar matrix (i.e., $\mathbf{M}\mathbf{M}^{\mathsf{H}} = \mu\mathbf{I}$), $p(\mathbf{W})$ only depends on the eigenvalues of $\mathbf{W}$. Thus $\mathbf{W}$ is unitarily invariant. In such a case, the distribution of $\mathbf{\Lambda}$ can be obtained from (2.27) by applying the limit in (2.7) and the property in (2.20), and it is given by

$$p(\mathbf{\Lambda}) = \frac{\pi_m^2 |\mathbf{\Lambda}|^n |\mathbf{F}| \mathcal{V}(\mathbf{\Lambda})}{\Gamma_m(m)\Gamma_m(n) \mathrm{e}^{\mu m + \mathsf{Tr}\{\mathbf{\Lambda}\}}} \tag{2.28}$$

  where $(\mathbf{F})_{ij} = \lambda_j^{-i} {}_0F_1(\,;n-i+1;\mu\lambda_j)$.

If $\mathbf{M}\mathbf{M}^{\mathsf{H}}$ is with lower L-rank, and the non-zero eigenvalues are $\mu_1 > \mu_2 > \cdots > \mu_L$, then the distribution of $\mathbf{\Lambda}$ is given by [30]

$$p(\mathbf{\Lambda}) = K(\mathbf{M})\mathcal{V}(\mathbf{\Lambda})\text{etr}(-\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}|\mathbf{T}| \tag{2.29}$$

where

$$K(\mathbf{M}) = \frac{\text{etr}(-\mathbf{M}\mathbf{M}^{\mathsf{H}})}{\Gamma_{m-L}(m-L)\Gamma_{m-L}(n-L)\mathcal{V}(\mathbf{M}\mathbf{M}^{\mathsf{H}})\prod_{i=1}^{L}\mu_i^{m-L}} \tag{2.30}$$

and $\mathbf{T}$ is a $m \times m$ matrix with (i,j)th entry as

$$(\mathbf{T})_{ij} = \begin{cases} {}_0F_1(n-m+1;\mu_j\lambda_i)/(n-m)! & i = 1,\ldots,m;\ j = 1,\ldots,L \\ \lambda_i^{m-j} & i = 1,\ldots,m;\ j = L+1,\ldots,m \end{cases} \tag{2.31}$$

Note that, since the Vandermonde determinant in (2.6) and pdf are positive by definition, here and in the following $|\mathbf{F}|$ represents the absolute value of the determinant of matrix $\mathbf{F}$. This avoids us to include in the provided results coefficients that account for the sign of determinants.

- When $n \leq p$, Beta-distributed $\mathbf{B}$ admits an eigendecomposition where the matrix of the eigenvectors is independent of the matrix of the eigenvalues [2, Lemma 8].

  - For $q \leq n$, the distribution of the q ordered non-zero eigenvalues of $\mathbf{B}$ is given by [2, eq. (13)]:

$$p(\mathbf{\Lambda}) = \frac{\pi_q^2\Gamma_q(p+q)|\mathbf{I}-\mathbf{\Lambda}|^{n-q}|\mathbf{\Lambda}|^{p-n}\mathcal{V}^2(\mathbf{\Lambda})}{\Gamma_q(n)\Gamma_q(p+q-n)\Gamma_q(q)}\,. \tag{2.32}$$

  - For $q > n$, $\mathbf{B}$ has n nonzero eigenvalues, whose ordered joint distribution is given by [2, eq. (12)]:

$$p(\mathbf{\Lambda}) = \frac{\pi_n^2\Gamma_n(p+q)|\mathbf{I}-\mathbf{\Lambda}|^{q-n}|\mathbf{\Lambda}|^{p-n}\mathcal{V}^2(\mathbf{\Lambda})}{\Gamma_n(n)\Gamma_n(p)\Gamma_n(q)}\,. \tag{2.33}$$

Due to the lack of the corresponding expression in the literature, herein we derive the expression of the marginal distribution of a single unordered eigenvalue of a $\mathcal{B}_n(p,q)$-distributed matrix, which will be needed in our subsequent derivations.

2.4 New Random Eigenvalue Results

In this section, we present the novel expressions for the eigenvalues distribution of various matrix-variate structures. Specifically, we derived the joint eigenvalues distribution of central and noncentral $\mathcal{F}$-matrix and the marginal eigenvalue distribution of Beta-distributed matrix. These new expressions will be utilized in this thesis to compute the performance measures of the multiple-antenna network.

Theorem 1. *Let* $\mathbf{H}_1$ *and* $\mathbf{H}_2$ *be, respectively, an* $m \times n$ *and an* $m \times p$ $(m \leq p)$ *Gaussian complex random matrix whose columns are independent, have zero mean, and covariance* $\mathbf{\Theta}_1$ *and* $\mathbf{\Theta}_2$*, respectively.*

- *For* $m \leq n$*, the* $m \times m$ *random matrix* $\mathbf{W} = (\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}\mathbf{H}_1\mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}$ *is a central* $\mathcal{F}$*-matrix [29]. When* $\mathbf{\Theta}_1$ *and* $\mathbf{\Theta}_2$ *are both scalar matrices,* $\mathbf{W}$ *is unitarily invariant and has a Beta type II distribution [36]. Specifically, when* $\mathbf{\Theta}_1\mathbf{\Theta}_2^{-1} = \omega\mathbf{I}$*, the distribution of its ordered eigenvalues is given by*

$$p(\mathit{\Lambda}) = \frac{\pi_m^2\Gamma_m(p+n)}{\omega^{mn}\Gamma_m(m)\Gamma_m(p)\Gamma_m(n)}\frac{\mathcal{V}^2(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}}{|\mathbf{I}+\mathbf{\Lambda}/\omega|^{p+n}}\,. \tag{2.34}$$

- *For* $m > n$*, the matrix* $\mathbf{W} = \mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1}\mathbf{H}_1$ *is unitarily invariant and the distribution of its ordered eigenvalues can be expressed as*

$$p(\mathbf{\Lambda}) = \frac{\pi_n^2\mathit{\Gamma}_m(p+n)|\mathbf{F}||\mathbf{\Omega}|^{-n}\mathcal{V}(\mathbf{\Lambda})|\mathbf{I}+\mathbf{\Lambda}|^{m-p-n-1}}{(p+n-m)!^{-n}\Gamma_n(p+n)\Gamma_m(p)\Gamma_n(n)\mathcal{V}(\mathbf{I}-\mathbf{\Omega}^{-1})} \tag{2.35}$$

  where $\mathbf{\Omega} = \mathbf{\Theta}_1^{1/2}\mathbf{\Theta}_2^{-1}\mathbf{\Theta}_1^{1/2}$*,* $(\mathbf{F})_{ij} = {}_1F_0(p+n-m+1;;(1-\omega_i^{-1})\lambda_j/(1+\lambda_j))$ *for* $i = 1,\ldots,m$*,* $j = 1,\ldots,n$*, and* $(\mathbf{F})_{ij} = (1-\omega_i^{-1})^{m-j}$ *for* $i = 1,\ldots,m$*,* $j = n+1,\ldots,m$*.*

Proof. The proof is given in Appendix A.1. □

Theorem 2. *Let* $\mathbf{H}_1$ *and* $\mathbf{H}_2$ *be, respectively, an* $m \times n$ *and an* $m \times p$ *Gaussian complex random matrix whose columns are independent and have covariance* $\mathbf{\Theta}$*. Let also* $\mathbb{E}[\mathbf{H}_1] = \mathbf{M}$ *and* $\mathbb{E}[\mathbf{H}_2] = \mathbf{0}$*.*

- *For* $m \leq n$*,* $\mathbf{M}\mathbf{M}^{\mathsf{H}} = \mu\mathbf{I}$ *and* $\mathbf{\Theta} = \theta\mathbf{I}$*, the non-central* $\mathcal{F}$*-matrix* $\mathbf{W} = (\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}\mathbf{H}_1\mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}$ *is unitarily invariant and the distribution of its eigenvalues is given by*

$$p(\mathbf{\Lambda}) = \frac{\pi_m^2\mathcal{V}(\mathbf{\Lambda})^2\Gamma_m(p+n){}_1F_1(p+n;n;\frac{\mu}{\theta}\mathbf{\Lambda}(\mathbf{I}+\mathbf{\Lambda})^{-1})}{\Gamma_m(m)\Gamma_m(n)\Gamma_m(p)\mathrm{e}^{\mu m/\theta}|\mathbf{\Lambda}|^{m-n}|\mathbf{I}+\mathbf{\Lambda}|^{p+n}} \tag{2.36}$$

- *For $m > n$, and $\mathbf{M}^{\mathsf{H}}\boldsymbol{\Theta}^{-1}\mathbf{M} = \omega\mathbf{I}$, the matrix $\mathbf{W} = \mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1}\mathbf{H}_1$ is unitarily invariant and the distribution of its eigenvalues is given by*

$$p(\boldsymbol{\Lambda}) = \frac{\pi_n^2\Gamma_n(p+n)\mathrm{e}^{-\omega n}}{\Gamma_n(n)\Gamma_n(m)\Gamma_n(p+n-m)}\frac{|\mathbf{F}||\boldsymbol{\Lambda}|^{m-n}\mathcal{V}(\boldsymbol{\Lambda})}{|\mathbf{I}+\boldsymbol{\Lambda}|^{p+1}} \tag{2.37}$$

where $(\mathbf{F})_{ij} = (\lambda_j/(1+\lambda_j))^{n-i}{}_1F_1(p+n-i+1; m-i+1; \omega\lambda_j/(1+\lambda_j))$.

Proof. The proof is provided in Appendix A.2. □

Theorem 3. *Given an $n \times n$ matrix $\mathbf{B} \sim \mathcal{B}_n(p,q)$,*

- *For $q \leq n$, the pdf of a single unordered eigenvalue of $\mathbf{B}$ is given by*

$$\begin{aligned} p(\lambda) &= \frac{\pi_q^2}{q\Gamma_q(q)}\frac{\Gamma_q(p+q)\Gamma(n-q+1)}{\Gamma_q(n)\Gamma_q(p+q-n)} \\ &\quad \cdot \sum_{i,j=1}^{n} \lambda^{p-n+i+j-2}(1-\lambda)^{n-q}\mathcal{D}_{ij} \end{aligned} \tag{2.38}$$

with $\mathcal{D}_{ij}$ being the (i,j)-cofactor of the $(n \times n)$ matrix $\mathbf{A}$ such that

$$(\mathbf{A})_{\ell k} = \frac{\Gamma(p-n+\ell+k-1)}{\Gamma(p+k-q+\ell)}\,. \tag{2.39}$$

- *For $q > n$, the pdf of a single unordered eigenvalue of $\mathbf{B}$ is given by*

$$\begin{aligned} p(\lambda) &= \frac{\pi_n^2}{\Gamma_n(n)}\frac{\Gamma_n(p+q)\Gamma(q-n+1)}{\Gamma_n(p)\Gamma_n(q)} \\ &\quad \cdot \sum_{i,j=1}^{n} \lambda^{(p-n+i+j-2)}(1-\lambda)^{q-n}\mathcal{D}_{ij} \end{aligned} \tag{2.40}$$

with $\mathcal{D}_{ij}$ being the (i,j)-cofactor of the $(n \times n)$ matrix $\mathbf{A}$ such that

$$(\mathbf{A})_{\ell k} = \frac{\Gamma(p-n+\ell+k-1)}{\Gamma(p+k+q-2n+\ell)}\,. \tag{2.41}$$

Proof. The proof is provided in Appendix A.3. □

Chapter 3

Closed-form Output Statistics of MIMO Block-Fading Channels

The information that can be transmitted through a wireless channel, with multiple-antenna equipped transmitter and receiver, is crucially influenced by the channel behavior as well as by the structure of the input signal. In this chapter, we characterize in closed form the probability density function of the output of MIMO block-fading channels, for an arbitrary SNR value. Our results provide compact expressions for such output statistics, paving the way to a more detailed analytical information-theoretic exploration of communications in presence of block fading. The analysis is carried out assuming two different structures for the input signal: the i.i.d. Gaussian distribution and a product form that has been proved to be optimal for non-coherent communication, i.e., in absence of any channel state information. When the channel is fed by an i.i.d. Gaussian input, we assume the Gramian of the channel matrix to be unitarily invariant and derive the output statistics in both the noise-limited and the interference-limited scenario, considering different fading distributions. When the product-form input is adopted, we provide the expressions of the output pdf as the relationship between the overall number of antennas and the fading coherence length varies. We also highlight the relation between our newly derived expressions and the results already available in the literature, and, for some cases, we numerically compute the mutual information, based on the proposed expression of the output statistics.

3.1 Introduction

The availability of an explicit statistical characterization of the output of a wireless channel, impaired by additive and multiplicative random disturbance, is of

paramount importance to communication- and information-theoretic purposes. Indeed, a closed-form expression for the output probability density function (pdf) is relevant for the evaluation of the ergodic mutual information between the input and the output signals of a randomly faded channel [19]. It also turns out to be crucial in the finite block-length regime, in order to characterize the information density of the communication at hand [37].

In spite of its importance, few explicit results are available in the literature for the output signal pdf in the case of MIMO block-independent fading channels. The works in [35, 38, 2] all focus on the case of block-Rayleigh fading. In these papers, the output statistics are derived under different assumptions on the relative values of the number of involved antennas and of the coherence length of the fading. The input distribution, too, plays a crucial role in the cited derivations. More specifically, in [35] the authors assume the input to be i.i.d. Gaussian and investigate the behavior of the output distribution as the fading coherence length varies from being quite short to very long, compared to the overall number of transmit and receive antennas. In both [38] and [2], instead, the input is assumed to be given by the product of a diagonal matrix (representing the power allocation over the transmit antennas) times an isotropically distributed matrix with unitary columns. The main difference between the two papers is in the assumption on the fading duration. Indeed, the first one focuses on the case where the coherence length of the Rayleigh fading is greater than the number of involved antennas; in this case, the high SNR-optimal power allocation matrix turns out to be a scaled version of the identity matrix [39]. The study in [2], instead, solves the problem of characterizing, again in the high-SNR regime, the optimal power allocation profile, assuming the fading coherence length to be shorter, compared to the number of involved antennas. In the latter case, indeed, the diagonal matrix of the power allocation is characterized by the eigenvalues of a matrix-variate Beta joint distribution of the entries [2].

In this chapter, we consider both the input models described above, and derive closed form expressions of the output pdf in presence of a multiple-antenna channel affected by additive noise and block-fading. In particular, in the case of i.i.d. Gaussian input, our procedure allows the derivation of a closed-form expression for the output statistics of channels with unitarily invariant fading law. Apart from the canonical i.i.d. Rayleigh fading, already treated in [35], this encompasses the Rician channel with scalar LoS matrix, whose analysis was previously limited to the evaluation of the fading number [40], and the LoS MIMO [41] with a certain amount of residual scattering. Also, we provide results for the Land Mobile Satellite (LMS) with scalar average power LoS matrix [42, Property I] and for the above cases of MIMO Rayleigh and Rician fading communications impaired by Rayleigh-faded co-channel interference [43]. We remark that the expressions of the output pdf that we derive hold for any arbitrary value of SNR.

3.2 Communication Model

We consider a single-user multiple-antenna communication system, with m and n denoting the number of receive and transmit antennas, respectively. Assuming block-memoryless fading with coherence length equal to b, the output can be described by the following linear relationship:

$$\mathbf{Y} = \sqrt{\gamma}\mathbf{H}\mathbf{X} + \mathbf{N} \tag{3.1}$$

where $\mathbf{Y}$ is the $m \times b$ output matrix, and $\mathbf{H}$ is the $m \times n$ complex random channel matrix whose entries represent the fading coefficients between each transmit and receive antenna. $\mathbf{N}$ is the $m \times b$ matrix of white Gaussian noise which is assumed to have i.i.d. complex Gaussian entries with zero mean and unitary variance. The normalized per-transmit antenna SNR is denoted by $\gamma = \mathrm{SNR}/n$, and $\mathbf{X}$ is the random complex $n \times b$ input matrix whose structure will be specified in the following sections. Moreover, for any positive integer n, we define

$$\gamma_n = \gamma^{n(n-1)/2} .$$

Note that the above communication model is adopted in all the following sections, except for Section 3.3.2 where we resort to a slightly different model explicitly accounting for interference.

3.3 Output Statistics with IID Gaussian Input

In this section, we analyse the case where the distribution of $\mathbf{X}$ is Gaussian i.i.d. and consider both the noise-limited and interference-limited scenarios. Note that, in the case under study, the average energy of the input signal is given by $\mathbb{E}[\mathsf{Tr}\{\mathbf{X}\mathbf{X}^{\mathsf{H}}\}] = nb$.

As for the communication channel, we focus our analysis on some classes of channel matrices whose Gramian $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$ is unitarily invariant. As shown in the following, this allows us to write the expression of the output pdf in terms of the distribution of the eigenvalues of the channel matrix. In particular, in both the noise-limited and the interference-limited case, we draw on the following results:

- for $m \leq n$, and for unitarily invariant $\mathbf{H}\mathbf{H}^{\mathsf{H}}$, the distribution of $\mathbf{Y}$ is given by [35, eq. (40) and (41)]

$$p(\mathbf{Y}) = \frac{\Gamma_m(m) K(\mathbf{Y})}{\pi_m \gamma_m} \int \frac{|\mathbf{E}||\mathbf{I} + \gamma\mathbf{\Lambda}|^{m-b-1}}{\mathcal{V}(\mathbf{\Lambda})} p(\mathbf{\Lambda})\, \mathrm{d}\mathbf{\Lambda} , \tag{3.2}$$

where $\mathbf{\Lambda}$ is an $m \times m$ diagonal matrix containing the eigenvalues of channel matrix $\mathbf{HH}^{\mathsf{H}}$, $(\mathbf{E})_{ij} = \mathrm{e}^{y_i c_j}$, and $c_j = \gamma\lambda_j/(1+\gamma\lambda_j)$, $j = 1,\ldots,m$. Moreover, $y_1,\ldots,y_m$ are the eigenvalues of $\mathbf{YY}^{\mathsf{H}}$ and

$$K(\mathbf{Y}) = \frac{\mathrm{e}^{-\|\mathbf{Y}\|^2}}{\pi^{mb}\mathcal{V}(\mathbf{YY}^{\mathsf{H}})}\,. \tag{3.3}$$

- for $m > n$, and for unitarily invariant $\mathbf{H}^{\mathsf{H}}\mathbf{H}$, the pdf of $\mathbf{Y}$ can be obtained by following the steps described in [35] and is given by

$$p(\mathbf{Y}) = \frac{\Gamma_n(m)K(\mathbf{Y})}{\pi_n\gamma^{n(m-n)}}\int \frac{|\widetilde{\mathbf{E}}||\mathbf{I}+\gamma\mathbf{\Lambda}|^{m-b-1}}{\mathcal{V}(\gamma\mathbf{\Lambda})|\mathbf{\Lambda}|^{m-n}}p(\mathbf{\Lambda})\,\mathrm{d}\mathbf{\Lambda} \tag{3.4}$$

where $\widetilde{\mathbf{E}}$ is an $m \times m$ matrix whose elements are given by $(\widetilde{\mathbf{E}})_{ij} = \mathrm{e}^{y_i c_j}$ for $1 \le j \le n$, and $(\widetilde{\mathbf{E}})_{ij} = y_i^{j-n-1}$ for $n+1 \le j \le m$. Note that in this case the matrix $\mathbf{HH}^{\mathsf{H}}$ is of reduced rank since it has $m-n$ zero eigenvalues. Thus, here $p(\mathbf{\Lambda})$ indicates the distribution of the n non-zero eigenvalues of $\mathbf{HH}^{\mathsf{H}}$ and $\mathbf{\Lambda}$ is an $n \times n$ diagonal matrix.

Proof. The proof is given in Appendix A.4. □

3.3.1 Noise-Limited

The output pdf of the uncorrelated Rayleigh-faded channel has been evaluated in [35]. For sake of completeness, we recall this result and present the corrected expression of the output pdf when $m > n$. Then, we extend the analysis to two other practically relevant fading models, namely, the Rician block-fading channel [44, 45] and Land Mobile Satellite (LMS) channel [42, 46].

Rayleigh Fading Channel

In the case of uncorrelated Rayleigh channel, the entries of $\mathbf{H}$ follow an i.i.d. zero-mean, unit-variance, complex Gaussian distribution.

- For $m \le n$, the distribution of the eigenvalues of $\mathbf{HH}^{\mathsf{H}}$ is given by (2.24). It follows that, by using (3.2) and the result in Appendix A.11, the distribution of $\mathbf{Y}$ can be written as [35, Proposition 2].

$$p(\mathbf{Y}) = \frac{\pi_m}{\gamma_m\Gamma_m(n)}K(\mathbf{Y})|\mathbf{Z}| \tag{3.5}$$

where the i,j-th entry of the $m \times m$ matrix $\mathbf{Z}$ is given by

$$(\mathbf{Z})_{ij} = \int_0^\infty \exp\left(\frac{y_i\gamma x}{1+\gamma x} - x\right)\frac{x^{n-m+j-1}}{(1+\gamma x)^{b+1-m}}\,\mathrm{d}x\,.$$

- For $m > n$, the distribution of the eigenvalues of channel matrix $\mathbf{H}^{\mathsf{H}}\mathbf{H}$ is given by (2.25). By applying (3.4) and the result in Appendix A.11, the output pdf is given by

$$p(\mathbf{Y}) = \frac{\pi_n}{\Gamma_n(n)\gamma^{n(m-n)}} K(\mathbf{Y})|\mathbf{Z}|\,. \tag{3.6}$$

  Note that the expression above differs from the one presented in [35, Proposition 2] in the term $\gamma^{n(m-n)}$, which appears at the denominator. The i,j-th entry of the $m \times m$ matrix $\mathbf{Z}$ can be written as

$$(\mathbf{Z})_{ij} = \int_0^\infty \exp\left(\frac{y_i\gamma x}{1+\gamma x} - x\right) \frac{(x/\gamma)^{j-1}}{(1+\gamma x)^{b+1-m}}\,\mathrm{d}x\,,$$

  for $1 \le i \le m$,$1 \le j \le n$ and $(\mathbf{Z})_{ij} = y_i^{j-n-1}$, for $1 \le i \le m$,$n+1 \le j \le m$.

Rician Channel

The Rician channel is traditionally modeled as a superposition of a scattered plus a LoS component, i.e.,

$$\mathbf{H} = \sqrt{\frac{\kappa}{\kappa+1}}\bar{\mathbf{H}} + \sqrt{\frac{1}{\kappa+1}}\widetilde{\mathbf{H}}\,. \tag{3.7}$$

In (3.7), κ is the Rician factor representing the ratio of the average power of the unfaded channel component to the faded channel component, the entries of $\widetilde{\mathbf{H}}$ are independent, zero-mean unit-variance complex Gaussian, and $\bar{\mathbf{H}}$ is a deterministic matrix representing the LoS component.

Specifically, for $m \le n$, we consider the special case $\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathsf{H}} = h\mathbf{I}$ (for $m > n$ we assume $\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}} = h\mathbf{I}$), where h is a positive parameter. This assumption reflects two main settings: the scalar LoS channel, introduced in [40] and therein already analyzed in the high-SNR regime, and the LoS MIMO with residual scattering [41]. Both models assume the LoS matrix to have high (full) rank. The one in [41] is suitable for MIMO backhaul links where antenna spacing is carefully designed and transmit-receive distance is fixed. Our model can be thought of as a Gaussian perturbation, with small variance, of the one in [41]. The model in [40], although being a sub-case of the one in [41] from the pure mathematical viewpoint, has played a major role in the early characterization of MIMO Rician channels, due to the amenability of diagonal [47] (and, in particular, scalar) non-centrality matrices for the derivation of the capacity-achieving input law.

Under the aforementioned assumption, the Gramian of the matrix $\mathbf{H}$ is unitarily invariant, since (2.27) now is solely dependent on $\mathbf{\Lambda}$. Therefore, the pdf of the output can be expressed as in the following proposition.

Proposition 1. *Given a channel as in* (3.1) *and* (3.7)*, with i.i.d. Gaussian input and Rician block-fading,*

- *for* $m \leq n$*, and* $\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathsf{H}} = h\mathbf{I}$*, the pdf of its output can be written as*

$$p(\mathbf{Y}) = \frac{\pi_m (1+\kappa)^{mn}}{\gamma_m \Gamma_m(n) \mathrm{e}^{\kappa h m}} K(\mathbf{Y}) |\mathbf{Z}| \,, \tag{3.8}$$

  where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \gamma x/(1+\gamma x)} {}_0F_1(\,;n-j+1;\hat{x})\,\mathrm{d}x}{\mathrm{e}^{(1+\kappa)x} x^{j-n}(1+\gamma x)^{b-m+1}}$$

- *for* $m > n$*, and* $\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}} = h\mathbf{I}$*, the following result holds*

$$p(\mathbf{Y}) = \frac{\pi_n (1+\kappa)^{nm}}{\gamma_n \gamma^{n(m-n)} \Gamma_n(n) \mathrm{e}^{\kappa h n}} K(\mathbf{Y}) |\mathbf{Z}| \tag{3.9}$$

  where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \gamma x/(1+\gamma x)} {}_0F_1(\,;m-j+1;\hat{x})\,\mathrm{d}x}{\mathrm{e}^{(1+\kappa)x} x^{j-n}(1+\gamma x)^{b-m+1}} \,,$$

  for $1 \leq i \leq m$,$1 \leq j \leq n$ *and* $(\mathbf{Z})_{ij} = y_i^{j-n-1}$*, for* $1 \leq i \leq m$,$n+1 \leq j \leq m$,

with $\hat{x} = \kappa(1+\kappa)hx$*.*

Proof. The proof is given in Appendix A.5. □

Land Mobile Satellite Communication

The Land Mobile Satellite (LMS) MIMO channel can be viewed as a non-central channel with random mean. Thus, the channel matrix model can be described as

$$\mathbf{H} = \bar{\mathbf{H}} + \widetilde{\mathbf{H}} \tag{3.10}$$

where the entries of $\widetilde{\mathbf{H}}$ are independent, zero-mean unit-variance complex Gaussian and $\bar{\mathbf{H}}$ is a random matrix. As shown in [42], in a LMS channel the matrix $\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathsf{H}}$ follows a matrix-variate $\Gamma(\alpha,\mathbf{\Omega})$ distribution [48] where α plays the role of a shape parameter, while $\mathbf{\Omega}$ is a scale parameter. Indeed, α can be viewed as a generalized number of degrees of freedom of the non-centrality parameter, while $\mathbf{\Omega}$ is related to the average power of the random LoS component, as discussed in detail in [42]. Assuming $\mathbf{\Omega} = \omega\mathbf{I}$, $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is unitarily invariant, as shown in [42, Property 1].

Under this assumption, the expression of the output pdf can be expressed as in the following proposition.

Proposition 2. *Given an LMS MIMO channel as in* (3.10) *with* $\mathbf{\Omega} = \omega\mathbf{I}$*,*

- *for $m \leq n$, the pdf of its output can be written as*

$$p(\mathbf{Y}) = \frac{\pi_m}{\gamma_m \Gamma_m(n)(1+1/\omega)^{m\alpha}} K(\mathbf{Y})|\mathbf{Z}|\,, \tag{3.11}$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \frac{\gamma x}{1+\gamma x}}\, {}_1F_1\left(\alpha - j + 1; n - j + 1; \frac{x}{1+\omega}\right)}{\mathrm{e}^x x^{j-n}(1+\gamma x)^{b-m+1}}\,\mathrm{d}x\,;$$

- *for $m > n$, the output pdf is given by:*

$$p(\mathbf{Y}) = \frac{\pi_n}{\gamma_n \gamma^{n(m-n)} \Gamma_n(m)(1+1/\omega)^{n\alpha}} K(\mathbf{Y})|\mathbf{Z}| \tag{3.12}$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \frac{\gamma x}{1+\gamma x}}\, {}_1F_1\left(\alpha - j + 1; m - j + 1; \frac{x}{1+\omega}\right)}{\mathrm{e}^x x^{j-n}(1+\gamma x)^{b-m+1}}\,\mathrm{d}x\,,$$

for $1 \leq i \leq m$,$1 \leq j \leq n$ and $(\mathbf{Z})_{ij} = y_i^{j-n-1}$, for $1 \leq i \leq m$,$n+1 \leq j \leq m$.

Proof. The proof is given in Appendix A.6. □

3.3.2 Interference Limited

We now consider the case where the main impairment to communication is represented by the co-channel interference. In particular, each interferer is seen from the direct link receiver under its own random channel, which we assume to be affected by Rayleigh fading, again with block-length b. We assume that there are L active interferers in the network, each equipped, for homogeneity, with the same number of antennas, n, as the transmitter of the useful signal. We evaluate the output pdf when a whitening filter is applied to the received signal and we consider two channel models. In the former, the desired signal undergoes Rayleigh fading; in the latter, the direct link is affected by Rician fading, i.e., we assume the existence of an LoS path between the useful transmitter and its intended receiver.

The received signal can be modeled as

$$\widetilde{\mathbf{Y}} = \sqrt{\gamma}\mathbf{H}_s\mathbf{X} + \mathbf{W} \tag{3.13}$$

where

$$\mathbf{W} = \sum_{\ell=1}^{L} \widehat{\mathbf{H}}_\ell \widehat{\mathbf{X}}_\ell$$

represents the interference. Specifically, the $m \times n$ matrix $\widehat{\mathbf{H}}_\ell$ models the channel connecting the ℓ-th interferer with the receiver, while the $n \times b$ matrix $\widehat{\mathbf{X}}_\ell$ represents

the signal transmitted by the ℓ-th interferer, $\ell = 1,\ldots,L$. The interference can be rewritten as $\mathbf{W} = \widehat{\mathbf{H}}\widehat{\mathbf{X}}$ where $\widehat{\mathbf{H}} = [\widehat{\mathbf{H}}_1,\ldots,\widehat{\mathbf{H}}_L]$ is an $m \times Ln$ matrix and $\widehat{\mathbf{X}} = [\widehat{\mathbf{X}}_1{}^{\mathsf{H}},\ldots,\widehat{\mathbf{X}}_L{}^{\mathsf{H}}]^{\mathsf{H}}$ is of size $Ln \times b$. By assuming that the entries of $\widehat{\mathbf{X}}$ are i.i.d. complex Gaussian with zero mean and unit variance, the covariance of the interference, conditioned on the knowledge of the composite channel matrix $\widehat{\mathbf{H}}$, is given by

$$\mathbf{R} = \mathbb{E}[\mathbf{W}\mathbf{W}^{\mathsf{H}}|\widehat{\mathbf{H}}] = \widehat{\mathbf{H}}\mathbb{E}[\widehat{\mathbf{X}}\widehat{\mathbf{X}}^{\mathsf{H}}]\widehat{\mathbf{H}}^{\mathsf{H}} = b\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\,.$$

We apply to the received signal $\widetilde{\mathbf{Y}}$ the whitening filter $\mathbf{B} = \sqrt{b}\mathbf{R}^{-1/2}$ and obtain

$$\begin{aligned}\mathbf{Y} &= \mathbf{B}\widetilde{\mathbf{Y}}\\ &= \sqrt{b}\mathbf{R}^{-1/2}\widetilde{\mathbf{Y}}\\ &= \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2}(\sqrt{\gamma}\mathbf{H}_s\mathbf{X}+\mathbf{W})\\ &= \sqrt{\gamma}\mathbf{H}\mathbf{X}+\mathbf{N}\end{aligned} \tag{3.14}$$

where $\mathbf{H} = \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2}\mathbf{H}_s$ and $\mathbf{N} = \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2}\mathbf{W}$. Clearly, $\mathbb{E}[\mathbf{N}\mathbf{N}^{\mathsf{H}}|\widehat{\mathbf{H}}] = b\mathbf{I}$. In the following, we provide the pdf of $\mathbf{Y}$.

Rayleigh Fading Channel

Proposition 3. *We consider the interference-limited channel described by* (3.13)*, with L active interferers, i.i.d. Gaussian input and Rayleigh fading. If $\mathbf{H}_s\mathbf{H}_s{}^{\mathsf{H}} \sim \mathcal{W}_m(n,\mathbf{\Theta_s})$ and $\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}} \sim \mathcal{W}_m(Ln,\widehat{\mathbf{\Theta}})$, then we have the following results.*

- *For $m \leq n$, due to mathematical constraints, we only analyse the case of spatially uncorrelated receiving antennas, i.e., $\mathbf{\Theta_s} = \theta_\mathbf{s}\mathbf{I}$ and $\widehat{\mathbf{\Theta}} = \hat{\theta}\mathbf{I}$. Then, the pdf of $\mathbf{Y}$ can be written as*

$$p(\mathbf{Y}) = \frac{\pi_m\Gamma_m(Ln+n)}{\gamma_m\omega^{mn}\Gamma_m(Ln)\Gamma_m(n)}K(\mathbf{Y})|\mathbf{Z}| \tag{3.15}$$

where $\omega = \theta_s/\hat{\theta}$ and

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i\frac{\gamma x}{1+\gamma x}}x^{n-j}}{(1+\gamma x)^{b-m+1}(1+x/\omega)^{Ln+n}}\,\mathrm{d}x\,.$$

This result is obtained by substituting (2.34) *in* (3.2) *and by exploiting the result in Appendix A.11.*

- *For $m > n$, the pdf of $\mathbf{Y}$ is given by:*

$$p(\mathbf{Y}) = \frac{\pi_n (Ln+n-m)!^m \Gamma_n(m)\Gamma_m(Ln+n)K(\mathbf{Y})}{\Gamma_n(Ln+n)\Gamma_n(n)\Gamma_m(Ln)\gamma_n\gamma^{n(m-n)}} \cdot \frac{|\mathbf{\Omega}|^{m-n-1}}{\mathcal{V}(\mathbf{\Omega})} \int \frac{|\widetilde{\mathbf{E}}||\mathbf{I}+\gamma\mathbf{\Lambda}|^{m-b-1}|\mathbf{F}|\,\mathrm{d}\mathbf{\Lambda}}{|\mathbf{\Lambda}|^{m-n}|\mathbf{I}+\mathbf{\Lambda}|^{Ln+n-m+1}} \tag{3.16}$$

where $\mathbf{\Omega} = \mathbf{\Theta}_1^{1/2}\mathbf{\Theta}_2^{-1}\mathbf{\Theta}_1^{1/2}$ and the matrices $\widetilde{\mathbf{E}}$ and $\mathbf{F}$ have been defined below (3.4) *and* (2.35)*, respectively. This result is obtained by substituting* (2.35) *in* (3.4)*. However, we cannot solve the integral by applying the result in Appendix A.11 directly. Indeed, although matrices $\widetilde{\mathbf{E}}$ and $\mathbf{F}$ are both of size $m \times m$, a portion of their columns and rows is composed of constant terms. Thus, we need to resort to the property of the determinant of block matrices, in order to obtain $n \times n$ blocks to which the result in Appendix A.11 can be applied. We skip the details of this procedure due to the cumbersome expressions that are involved.*

Rician Fading Channel

Proposition 4. *We consider the interference-limited channel described by* (3.13)*, with L active interferers, i.i.d. Gaussian input, Rician faded useful signal and Rayleigh fading affecting the interfering links. For a Rician channel, matrix $\mathbf{H}_s$ can be written as in* (3.7)

$$\mathbf{H}_s = \sqrt{\frac{\kappa}{\kappa+1}}\bar{\mathbf{H}}_s + \sqrt{\frac{1}{\kappa+1}}\widetilde{\mathbf{H}}_s$$

where κ is the Rician factor, $\bar{\mathbf{H}}_s$ is deterministic, and $\widetilde{\mathbf{H}}_s$ is complex Gaussian with independent colums whose covariance is $\mathbf{\Theta}$. According to our assumptions on LoS links made in Section 3.3.1, we have:

- *for $m \leq n$, setting $\mathbf{\Theta}_s = \widehat{\mathbf{\Theta}} = \theta\mathbf{I}$ and $\bar{\mathbf{H}}_s\bar{\mathbf{H}}_s^{\mathsf{H}} = h\mathbf{I}$,*

$$p(\mathbf{Y}) = \frac{\pi_m \Gamma_m(Ln+n)\mathrm{e}^{-h\kappa m/\theta}}{\gamma_m\Gamma_m(n)\Gamma_m(Ln)\tilde{\kappa}^{-mn}}K(\mathbf{Y})|\mathbf{Z}| \tag{3.17}$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i\frac{\gamma x}{1+\gamma x}}{}_1F_1(\tilde{L}+j; n-m+j; h\kappa\tilde{\kappa}\tilde{x}/\theta)\,\mathrm{d}x}{(1+\gamma x)^{b-m+1}(1+\tilde{\kappa}x)^{\tilde{L}+1}x^{m-n}\tilde{x}^{1-j}} \tag{3.18}$$

with $\tilde{\kappa} = 1+\kappa$, $\tilde{L} = Ln+n-m$, and $\tilde{x} = x/(1+\tilde{\kappa}x)$

- *for $m > n$, and $\bar{\mathbf{H}}_s{}^{\mathsf{H}}\boldsymbol{\Theta}^{-1}\bar{\mathbf{H}}_s = h\mathbf{I}$,*

$$p(\mathbf{Y}) = \frac{\pi_n \Gamma_n(Ln+n)\tilde{\kappa}^{nm}\mathrm{e}^{-h\kappa n}}{\gamma_n \gamma^{n(m-n)}\Gamma_n(n)\Gamma_n(Ln+n-m)} K(\mathbf{Y})|\mathbf{Z}| \tag{3.19}$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{{}_1F_1(Ln+n-j+1; m-j+1; h\kappa\tilde{\kappa}\tilde{x})\,\mathrm{d}x}{\mathrm{e}^{\frac{-y_i\gamma x}{1+\gamma x}}(1+\tilde{\kappa}x)^{Ln+1}\tilde{x}^{j-n}(1+\gamma x)^{b-m+1}},$$

for $1 \le i \le m$,$1 \le j \le n$, and $(\mathbf{Z})_{ij} = y_i^{j-n-1}$, for $1 \le i \le m$,$n+1 \le j \le m$; with $\tilde{\kappa} = 1+\kappa$ and $\tilde{x} = x/(1+\tilde{\kappa}x)$.

Proof. The proof is given in Appendix A.7. □

Note that also in this case mathematical issues made the analysis only possible for uncorrelated receivers.

3.3.3 Exploitation of The Analytical Results

The mutual information between the channel input, $\mathbf{X}$, and the channel output, $\mathbf{Y}$, normalized to the fading coherence length, can be expressed as:

$$\mathcal{I} = \frac{1}{b}\left[h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X})\right] \tag{3.20}$$

where $h(\mathbf{Y}) = \mathbb{E}[-\log p(\mathbf{Y})]$ and $h(\mathbf{Y}|\mathbf{X}) = \mathbb{E}[-\log p(\mathbf{Y}|\mathbf{X})]$. Once the pdf of the channel output, $p(\mathbf{Y})$, is obtained, it can be used to evaluate its differential entropy, $h(\mathbf{Y})$. For Rayleigh and Gaussian channels with identity covariance matrix, considering that $\mathbf{X}$ is given, the output $\mathbf{Y}$ is complex Gaussian and its rows are i.i.d. Hence, in order to derive the conditional differential entropy $h(\mathbf{Y}|\mathbf{X})$, we can compute its value for an arbitrary row of $\mathbf{Y}$ and then scale it by the number of rows of $\mathbf{Y}$ [35].

In [35], the mutual information has been computed in presence of Rayleigh channel and i.i.d. Gaussian input, for $m \le n$. In the following, we provide three examples of mutual information computation. First, we address the case of noise-limited Rayleigh channel with $m > n$ and, then, the noise-limited Rician channel, both with $m \le n$ and $m > n$.

In the case of Rayleigh channel, the conditional differential entropy is obtained using [35, eq. (4)], while the unconditional differential entropy is evaluated using (3.5) or (3.6) depending on the relationship between m and n. Fig. 3.1 shows the mutual information as a function of the SNR, with $b = 6$,10, $m = 2$ and $n = 1$, when no channel state information (CSI) is available and in the case of perfect CSI at the receiver. The latter is obtained by computing [35, eq. (10)]. The results

confirm the intuition, as well as previous analysis [38, 39]: the higher the SNR and the value of b, the better the performance, while the lack of CSI causes a noticeable degradation.

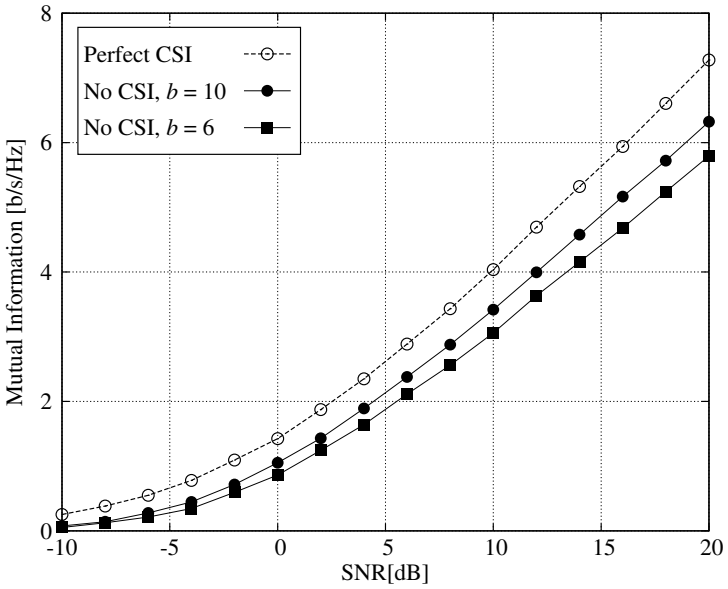

Figure 3.1: Mutual information vs. SNR in Rayleigh channel: comparison between the case where no CSI is available (solid line) and the case of perfect CSI at the receiver (dashed line), with $b = 6{,}10$, $m = 2$ and $n = 1$.

For the Rician channel, the expression of the channel matrix is given by (3.7). By adopting again the method in [35], the differential entropy of the output conditioned on the input signal can be computed. Let us denote by $\mathbf{y}$ an arbitrary row of $\mathbf{Y}$; then, using [35, eq. (31)] and considering the translation-invariant property of differential entropy, we can write the mutual information when the receiver does not have any knowledge of the non-LoS component:

$$h(\mathbf{y}|\mathbf{X}) = h(\mathbf{y}^{\mathsf{H}}|\mathbf{X}) = \mathbb{E}\left[\log_2\left((\pi \mathrm{e})^b \left|\mathbf{I} + \frac{\Gamma \mathbf{X}^{\mathsf{H}}\mathbf{X}}{1+\kappa}\right|\right)\right] \tag{3.21}$$

with the expectation being over the distribution of $\mathbf{X}$. The above expression can be conveniently computed resorting to [35, eq. (4)]. The unconditional differential entropy of the output is derived through (3.8) and (3.9).

Fig. 3.2 shows the mutual information as a function of the SNR, with $b = 6$, $m = 2$ and $n = 2$. Rician factors are set to $\kappa = 1$ and $\kappa = 10$. The plot depicts the mutual information in the two cases where the receiver has knowledge of the non-LoS component [35, eq. (10)] and where it does not (3.21). The deterministic

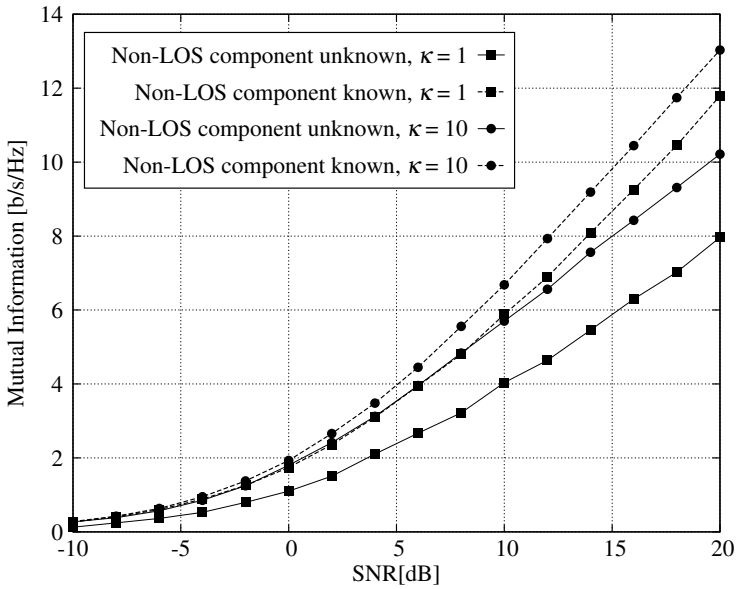

Figure 3.2: Mutual information vs. SNR in Rician channel: comparison between the case where the receiver does not have any knowledge on the non-LoS component (solid line) and when such knowledge is available (dashed line), for $b = 6$, $n = 2$, $m = 2$ and $\kappa = 1$,10.

channel matrix in (3.7) is set as follows:

$$\bar{\mathbf{H}} = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}.$$

In Fig. 3.2, the relative gap between the achievable mutual information in the two scenarios with $\kappa = 1$ is more evident than for $\kappa = 10$, since the higher the Rician factor, the higher the amount of information on the LoS component, which is known at the receiver. This is also compliant with the monotonicity results in [47].

Finally, Fig. 3.3 shows the mutual information for the two scenarios above, in the case of $m > n$, namely, $m = 2$, $n = 1$, and $b = 6$. The Rician factor is set to $\kappa = 1$ and $\kappa = 5$. In this scenario, the deterministic channel matrix is set to $\bar{\mathbf{H}} = \left[\sqrt{3/2}, 1/\sqrt{2}\right]^H$. Similar observations to those above hold. However, comparing Fig. 3.2 to Fig. 3.3, we notice that, as expected, the reduction in the number of antennas at the transmitter leads to severe performance degradation.

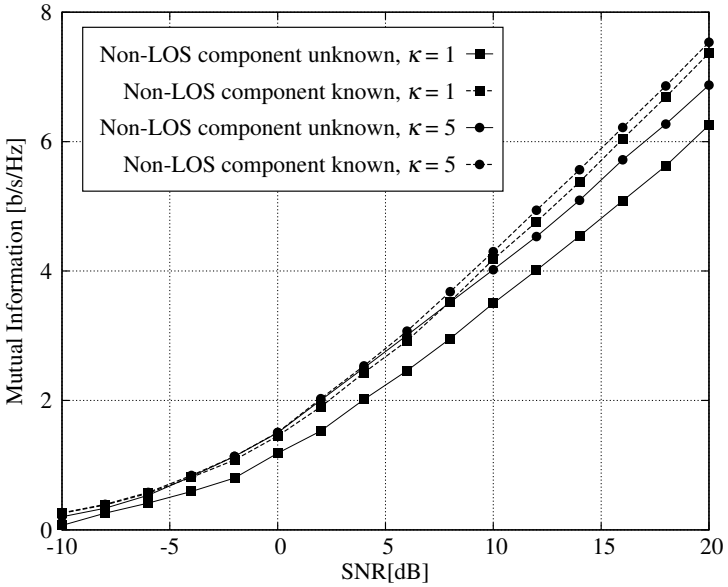

Figure 3.3: Mutual information vs. SNR in Rician channel: comparison between the cases where knowledge of the non-LoS component is not available at the receiver (solid line) and when it is (dashed line). $b = 6$, $m = 2$, $n = 1$ and $\kappa = 1{,}5$.

3.4 Output Statistical Characterization with Product Input Form

As in [35, 39, 8, 49], we assume total lack of CSI at both the ends of the wireless link. This case is of particular interest for the energy efficiency of the communication, as the availability of CSI would imply a high energy and time consumption at both the transmitter and the receiver. Under this assumption, in the high-SNR regime, the capacity-achieving input matrix $\mathbf{X}$ is proven to have a product structure [8, theorem 2] and can be written as

$$\mathbf{X} = \sqrt{c}\mathbf{D}^{1/2}\mathbf{\Phi} \tag{3.22}$$

where c is a normalizing constant and $\mathbf{D}$ is a real random $n \times n$ diagonal matrix, which is positive definite with probability 1. The entries of $\mathbf{D}$ represent the amount of transmit power allocated to each of the n transmit antennas, while $\mathbf{\Phi} \in \mathcal{S}(\mathbf{n},\mathbf{b})$ represents the beamforming $n \times b$ matrix. In order to be consistent with the definition of SNR, we impose the constraint on the average input energy $\mathbb{E}[\mathsf{Tr}\{\mathbf{X}\mathbf{X}^{\mathsf{H}}\}] = nb$. It follows that, for our specific input structure, the normalizing constant is given by:

$$c = \frac{nb}{\mathbb{E}[\mathsf{Tr}\{\mathbf{D}\}]}\,. \tag{3.23}$$

In [2, Lemma 10], it is proven that without CSI, in Rayleigh block-fading channels, the optimal power allocation at the transmitter depends on the relationship

between the coherence length, b, and the total number of antennas at both the transmitter and receiver. Specifically,

- If $b \geq m+n$, all diagonal entries of $\mathbf{D}$ are almost surely equal to 1. This case corresponds to the conventional unitary space-time modulation (USTM) [8], where $\mathbf{D} = \mathbf{I}$ and $c = b$;
- If $b < m+n$, the optimal input is $\mathbf{D} \sim \mathcal{B}_n(b-n, m+n-b)$, which is referred to as Beta-variate space-time modulation (BSTM) [2]. This scenario allows the analysis of an uplink massive-MIMO system, with $m \geq n$ and even $m \gg n$, which is relevant in the next-generation cellular setting.

3.4.1 Case $b \geq m+n$

As mentioned above, when $b \geq m+n$, the optimal power allocation over the transmitter antennas is given by a diagonal matrix, $\mathbf{D}$, with entries almost surely equal to 1. Under these assumptions, the following results hold.

Proposition 5. *Consider a channel as in* (3.1)*, affected by i.i.d. block-Rayleigh fading and with input given by* (3.22)*. Let* $\mathbf{\Delta} = \gamma c \mathbf{D}(\mathbf{I} + \gamma c \mathbf{D})^{-1} = diag(\delta_1, \ldots, \delta_n)$*, with* δ_i*'s being distinct values. Then,*

- *for* $m \leq n$*, the pdf of its matrix-variate output, conditioned on* $\mathbf{D}$ *and for* $n \leq b$*, can be expressed as*

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\Gamma_m(b) K(\mathbf{Y}) |\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{m-n} |\mathbf{G}|}{\pi_m (b-n)!^m \mathcal{V}(\mathbf{\Delta}) |\mathbf{I} + \gamma c \mathbf{D}|^m} \tag{3.24}$$

  where for $j = 1, \ldots, n$

$$(\mathbf{G})_{ij} = \begin{cases} {}_1F_1(1; b-n+1; y_i \delta_j) & i = 1, \ldots, m \\ \delta_j^{n-i} & i = m+1, \ldots, n \end{cases}$$

- *for* $m > n$*, the conditioned output pdf becomes*

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\Gamma_n(b) K(\mathbf{Y}) |\mathbf{\Delta}|^{n-m} |\mathbf{G}|}{\pi_n (b-m)!^n \mathcal{V}(\mathbf{\Delta}) |\mathbf{I} + \gamma c \mathbf{D}|^m} \tag{3.25}$$

  where i=$1, \ldots, m$*, are given by*

$$(\mathbf{G})_{ij} = \begin{cases} {}_1F_1(1; b-m+1; y_i \delta_j) & j = 1, \ldots, n \\ y_i^{m-j} & j = n+1, \ldots, m\,. \end{cases}$$

Proof. The proof is given in Appendix A.8. □

Case $\mathbf{D} = \mathbf{I}$

The expressions of $p(\mathbf{Y}|\mathbf{D})$ in (3.24) and (3.25) hold provided that the diagonal elements of $\mathbf{D}$ are distinct. Thus, in general, the unconditional pdf of $\mathbf{Y}$ can be derived by integrating $p(\mathbf{Y}|\mathbf{D})$ over the distribution of $\mathbf{D}$. In this section, however, we focus on a particular power allocation matrix, $\mathbf{D} = \mathbf{I}$, and, by (3.23), we consider $c = b$. Note that, in this case the elements of $\mathbf{D}$ are not distinct, and expressions (3.24) and (3.25) cannot be directly evaluated. Indeed, $|\mathbf{G}| = 0$ and $\mathcal{V}(\mathbf{\Delta}) = \mathbf{0}$, and again a limit procedure must be applied.

We first observe that, for $\mathbf{D} = \mathbf{I}$ and $c = b$, we have $\mathbf{\Delta} = \gamma\mathbf{b}\mathbf{D}(\mathbf{I} + \gamma\mathbf{b}\mathbf{D})^{-\mathbf{1}} = \bar{\delta}\mathbf{I}$ where $\bar{\delta} = \frac{\gamma b}{1+\gamma b}$.

- For $m \leq n$, we apply the limit in (2.7) to the ratio $|\mathbf{G}|/\mathcal{V}(\mathbf{\Delta})$ in (3.24) and, after some algebra, obtain

$$\lim_{\mathbf{\Delta}\to\bar{\delta}\mathbf{I}} \frac{|\mathbf{G}|}{\mathcal{V}(\mathbf{\Delta})} = \frac{\pi_n \Gamma_m(n)(b-n)!^m}{\Gamma_n(n)\Gamma_m(b)} |\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{n-m} |\widehat{\mathbf{G}}|$$

  where $\widehat{\mathbf{G}}$ is an $m \times m$ matrix whose elements are given by $(\widehat{\mathbf{G}})_{ij} = y_i^{m-j}\, {}_1F_1(n-j+1; b-j+1; y_i\bar{\delta})$, $i = 1,\ldots,m$, $j = 1,\ldots,m$. By recalling (3.24), the distribution of $\mathbf{Y}$ is then given by

$$p(\mathbf{Y}) = \frac{\pi_n \Gamma_m(n)}{\pi_m \Gamma_n(n)} \frac{K(\mathbf{Y})|\widehat{\mathbf{G}}|}{(1+\gamma b)^{nm}}. \tag{3.26}$$

- For $m > n$, we apply the limit in (2.7) to (3.25) and obtain

$$\lim_{\mathbf{\Delta}\to\bar{\delta}\mathbf{I}} \frac{|\mathbf{G}|}{\mathcal{V}(\mathbf{\Delta})} = \frac{\pi_n (b-m)!^n |\widehat{\mathbf{G}}|}{\Gamma_n(b-m+n)}$$

  where in this case

$$(\widehat{\mathbf{G}})_{ij} = y_i^{n-j}\, {}_1F_1(n-j+1; b-m+n-j+1; y_i\bar{\delta})$$

  for $i = 1,\ldots,m$,$j = 1,\ldots,n$, and $(\widehat{\mathbf{G}})_{ij} = y_i^{m-j}$ for $i = 1,\ldots,m$,$j = n+1,\ldots,m$.

  By recalling (3.25), it follows that

$$p(\mathbf{Y}) = \frac{\Gamma_n(b) K(\mathbf{Y}) \bar{\delta}^{n(n-m)} |\widehat{\mathbf{G}}|}{\Gamma_n(b-m+n)(1+\gamma b)^{nm}}. \tag{3.27}$$

We remark that, under the above assumptions, the output pdf also appears in [38]. The corresponding derivations provided therein involve Fourier integrals and Hankel matrices thus resulting in a slightly less compact form than ours.

3.4.2 A Massive MIMO Regime: $b < m + n$

Now, we consider the case of $b < m + n$; an instance of this scenario, by letting $m \gg n$, can adequately model the reverse link of the celebrated massive-MIMO channel [50]. In presence of uncorrelated block-Rayleigh fading, the high-SNR capacity-achieving input structure, as already mentioned, departs from the equal power allocation and is Beta distributed. We provide herein the output pdf for a block-fading channel fed by BSTM [2].

Proposition 6. *Given a channel as in (3.1), with* $\mathbf{X} = \sqrt{c}\mathbf{D}^{1/2}\mathbf{\Phi}$, $n \leq b$, *and* $\mathbf{D} \sim \mathcal{B}_n(b-n,n+m-b)$, *the pdf of its output can be written as*

$$p(\mathbf{Y}) = \frac{\pi_n \Gamma_n(b)\Gamma_n(m)(\gamma c)^{n(n-b)} K(\mathbf{Y})|\mathbf{F}_4||\mathbf{Z}|}{\gamma_n c^{n(n-1)/2}\Gamma_n(n)\Gamma_n(b-n)\Gamma_n(n+m-b)} \tag{3.28}$$

where $\mathbf{Z}$ *is an* $n \times n$ *matrix, whose generic entry is given by:*

$$\begin{aligned}(\mathbf{Z})_{ij} \;=\; & \int_0^1 \frac{(1-x)^{m-b}x^{i-1-n}}{(1+c\gamma x)^{m-b+1}} \\ & \cdot \left[\mathrm{e}^{y_j \frac{c\gamma x}{1+c\gamma x}} - \sum_{\ell,k=1}^{b-n} \frac{(\mathbf{F}_4^{-1})_{\ell k}}{y_j^{n+k-b}} \mathrm{e}^{y_{\ell+n}\frac{c\gamma x}{1+c\gamma x}} \right] \mathrm{d}x\end{aligned} \tag{3.29}$$

with $(\mathbf{F}_4)_{ij} = y_{n+i}^{b-n-j}$ $i,j = 1,\ldots,b-n$.

Proof. The proof is given in Appendix A.9. □

3.4.3 Exploitation of The Analytical Results

We now use the above results to compute the achievable mutual information in a massive MIMO case. In order to derive the output differential entropy conditioned on the input signal, $h(\mathbf{Y}|\mathbf{X})$, we exploit the analytic expression of the conditional pdf of the output, $p(\mathbf{Y}|\mathbf{X})$, obtained above.

Proposition 7. *Given a channel as in* (3.1), *the differential entropy of the output,* $\mathbf{Y}$, *conditioned on the channel input,* $\mathbf{X}$, *can be written as:*

$$\begin{aligned}h(\mathbf{Y}|\mathbf{X}) \;=\; & bm\log_2(\pi\mathrm{e}) + Km\sum_{i,j=1}^{n} a_{ij} \sum_{\ell=0}^{m-b} \frac{(-1)^\ell \binom{m-b}{\ell}}{s_{i,j,\ell}-1} \\ & \cdot \left[\log_2(1+c\gamma) - \frac{c\gamma\, {}_2F_1(1,s_{i,j,\ell};s_{i,j,\ell}+1;-\gamma)}{s_{i,j,\ell}\ln 2} \right]\end{aligned} \tag{3.30}$$

where K is a constant term, $s_{i,j,\ell} = b - 2n + i + j + \ell$, and a_{ij} is the (i,j)-cofactor of an $n \times n$ matrix $\mathbf{A}$ such that

$$\mathbf{A}_{\ell k} = \frac{\Gamma(b - 2n + \ell + k - 1)\Gamma(m - n + 1)}{\Gamma(b - 3n + m + \ell + k)} .$$

Proof. The proof is given in Appendix A.10. □

The mutual information obtained in a massive-MIMO-like case is shown in the following figures. Fig. 3.4 depicts the mutual information for $n = 1$, as the SNR varies and m grows up to very large values. The plot also compares our results (denoted by markers) are compared to the approximation given in [2] for the high SNR regime (dashed lines). The two sets of curves match very closely for any value of the parameters, as expected due to the tightness of [2, eq. (8)]. As m varies, all three curves have the same slope, as this has been proven to be insensitive to the number of receiving antennas in our setting [2, eq. (8)]. As expected, better performance is obtained as m increases. However, interestingly, Fig. 3.5 shows that a much higher improvement can be achieved as the fading coherence length and the number of antennas at the transmitter sightly increase while m is fixed to 10. In particular, by comparing the two plots, a limited gain in performance is obtained when m increases, while, as expected, the mutual information growth is significant when n is increased by 1.

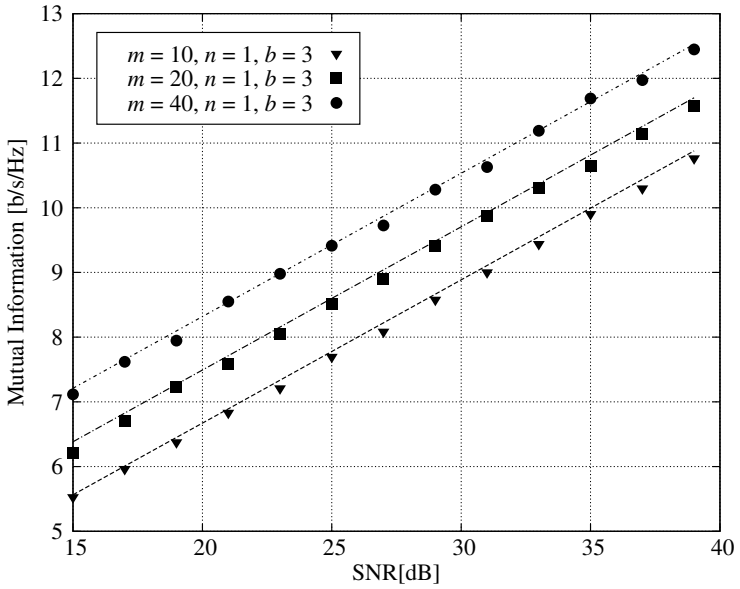

Figure 3.4: Mutual information vs. SNR in massive MIMO channel with BSTM: $b = 3$, $n = 1$ and different values of m. Our results (denoted by markers) are compared to the approximation in [2] (dashed lines).

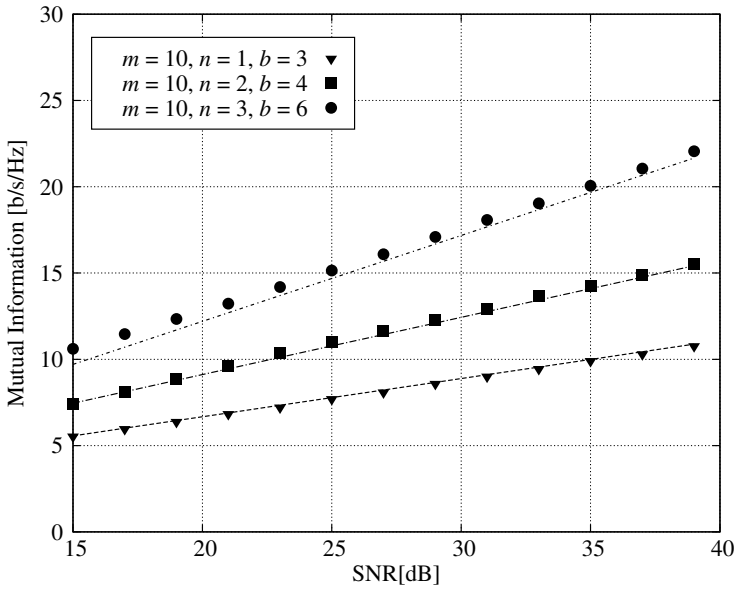

Figure 3.5: Mutual information vs. SNR in massive MIMO channel with BSTM: $m = 10$ and different values of b and n. Our results (denoted by markers) are compared to the approximation in [2] (dashed lines).

3.5 Conclusion

We obtained new, closed-form expressions for the probability density function of the output signal of a block-fading MIMO channel. By relying on recent results from the field of finite-dimensional random matrix theory, we provided results for the case of an i.i.d. Gaussian input under the assumption that the Gramian of the channel matrix is unitarily invariant. We addressed both the cases of Rayleigh and Rician fading. Furthermore, we derived the output probability density function in the case of product-form input. We particularized our newly derived expressions to those already available in the literature for the canonical case of uncorrelated Rayleigh fading, and we characterized the output signal behavior under different assumptions on the amplitude fading distribution.

Chapter 4

Characterization of Output Signals for MIMO Block-Fading Channels with Imperfect CSI

In this chapter, we provide an analytical characterization of the pdf of the output of a single-user, multiple-antenna communication link, impaired by block-fading. The analysis is carried out under the assumption of product structure for the input, assuming Gaussian noise, and an imperfect estimation of the channel available to the receiver. The model can be thought of as a perturbation of the case where the statistics of the channel are perfectly known. Specifically, the average singular values of the channel are given, while the channel singular vectors are assumed to be isotropically distributed on the unitary groups of dimensions given by the number of transmit and receive antennas. The channel estimate is affected by a Gaussian distributed error, which is modeled as a matrix with i.i.d. Gaussian entries of known covariance.

4.1 Introduction

The availability of an explicit statistical characterization for the output of a wireless channel is very relevant in both its communication- as well as its information-theoretic performance analysis. Few explicit results are available in the literature for the output signal pdf in the case of MIMO block-independent fading channels. Among these, there are [38, 35, 2], which all focus on the case of block-Rayleigh fading. In those papers, the output statistics are derived under different assumptions on the relative values of the number of involved antennas and of the coherence length of the fading. The input distribution, too, plays a crucial role in the cited derivations. More specifically, in [35] the authors assume the input to be i.i.d. Gaussian and

investigate the behavior of the output distribution as the fading coherence length varies from being quite shorter, compared to the overall number of transmit and receive antennas, up to a very long fading duration. In both [38] and [2], instead, the input is assumed to be given by the product of a diagonal matrix (representing the power allocation among the transmit antennas) times an isotropically distributed matrix with unitary columns. The main difference between the two papers is the assumption on the fading duration. Indeed, the first one focuses on the case where the coherence length of the Rayleigh fading is greater than the number of involved antennas; in this case, the optimal power allocation matrix turns out to be a scaled version of the identity matrix [39]. On the other hand, [2] solves the problem of characterizing, in the high SNR regime, the optimal power allocation profile, assuming the fading coherence length to be very short compared to the number of antennas in the wireless link. In this last case, indeed, the diagonal matrix of the power allocation is characterized by a Beta joint distribution of the entries.

In this chapter, we consider a single-user, multiple-antenna channel, affected by AWGN and block fading. Only partial CSI is available to the receiver, under the form of an imperfect estimation of the channel, affected by a Gaussian-distributed estimation error, as assumed, e.g. in [51] for the non block-fading case. In this scenario, the MIMO channel, as seen by the receiver can be adequately modeled as an estimated matrix, plus a matrix of uncorrelated, jointly Gaussian entries. Due to the AWGN assumption, this yields a non-central Gaussian behavior of the output signal, conditionally on the input structure and on the channel estimate. The main difference with respect to previous results is the model of the estimated channel matrix. This leads to a slightly more difficult algebra, and the desired statistics are to be evaluated relying on tools usually exploited in advanced analysis of non-central Wishart matrices (see e.g. [52, and references therein]). Some open mathematical issues in the evaluation of the output pdf as a function of the estimation quality are also reported and discussed.

4.2 System Model

We consider a single-user multiple-antenna communication, with m and n denoting the number of receive and, respectively, of transmit antennas. Assuming block-fading with block length b, the channel output can be described by the following linear relationship:

$$\mathbf{Y} = \sqrt{\gamma}\mathbf{H}\,\mathbf{X} + \mathbf{N}. \tag{4.1}$$

In (4.1), $\mathbf{Y}$ is the $m \times b$ output, $\mathbf{X}$ is the complex $n \times b$ input matrix, and $\mathbf{N}$ is the $m \times b$ matrix of additive complex circularly symmetric Gaussian noise. $\gamma = \mathrm{SNR}/n$ represents the normalized per-transmit antenna SNR.

$\mathbf{H}$ is a $m \times n$ complex channel matrix, whose entries represent the fading coefficients between each transmit and each receive antenna. Our assumption of partial CSI to the receiver implies that the channel matrix can be expressed as $\mathbf{H} = \widehat{\mathbf{H}} + \mathbf{E}$, with $\mathbf{E}$ the zero-mean matrix of the estimation errors, with i.i.d. entries [1] of variance σ_ϵ^2. The statistics of $\widehat{\mathbf{H}}$ will be specified, according to different assumptions in the following sections. Nevertheless, we will assume in general that the channel matrix admits a Singular Eigenvalue Decomposition, i.e., we can always write $\widehat{\mathbf{H}} = \mathbf{U}\widehat{\mathbf{\Sigma}}\mathbf{V}^{\mathsf{H}}$. $\mathbf{U}$ is an isotropic square unitary matrix of dimension m and $\mathbf{V}$ isotropic of dimension n.

The input matrix $\mathbf{X}$ is assumed to have a product structure, i.e.[2], $\mathbf{X} = \mathbf{D}^{1/2}\mathbf{\Phi}$, where $\mathbf{D}$ is a random, n-dimensional, diagonal matrix, which is positive definite w.p. 1. The diagonal entries of $\mathbf{D}$ are denoted by d_i's, $i = 1, \ldots, n$, and represent the amount of transmit power allocated to each of the n transmit antennas. $\mathbf{\Phi}$ is an $n \times b$ isotropic matrix, such that $\mathbf{\Phi}\mathbf{\Phi}^\dagger = \mathbf{I_n}$. We will refer to square isotropic matrices as Haar, and to rectangular isotropic matrices as Stiefel, respectively (see e.g. [38, 39]).

4.3 Statistical Characterization of The Channel Output

In this Section, we provide our main result, along with its detailed proof. We remark that the obtained results can be further particularized to two relevant cases, namely $\mathbf{D}$ proportional to a Beta-distributed matrix-variate and $\mathbf{D} = c\mathbf{I}$. Recall that $\mathbf{D} = c\mathbf{I}$ is optimal in the case of total absence of CSI as $b \geq n + m$, and is referred to as the Unitary Space Time Modulation (USTM) [38], while the Beta-distributed $\mathbf{D}$ provides optimal performance as $b < n + m$. In order to derive our results, we assume that $n \leq \min\{r,m\}$.

Proposition 8. *Given a channel as in (3.1), assume that an MMSE estimate of the channel matrix* $\mathbf{H}$ *is provided to the receiver, namely* $\widehat{\mathbf{H}}$ *has i.i.d. complex Gaussian entries with zero-mean and variance* $1-\sigma_\epsilon^2$*. The pdf of the channel output, conditionally on the input power allocation matrix* $\mathbf{D}$*, can be expressed as*

$$p(\mathbf{Y}|\mathbf{D}) = \frac{|\widetilde{\mathbf{F}}|e^{-\|\mathbf{Y}\|^2}|\mathbf{I}_n + \gamma\mathbf{D}|^{b-n-m}\prod_{i=b-n+1}^{b}\Gamma(i)}{(\gamma^n|\mathbf{D}|)^{b-n}\,\pi^{mb}\mathcal{V}(\mathbf{Y}^{\mathsf{H}}\mathbf{Y})\mathcal{V}(\widetilde{\mathbf{\Lambda}}^{-1})v_{\mathcal{S}}^{-1}}$$

[1] This model is a particular case of [53, Hypothesis II]. Indeed, it can be thought as being a permutation of the perfect receiver side information case. As the error variance $\sigma_\epsilon^2 \to 0$, this models the instantaneous receiver CSI.

[2] This input structure is relevant to the analysis since it is capacity-achieving in absence of CSI at both the ends of the link [39, Thm2].

where

$$v_{\mathcal{S}} = \frac{\tilde{v}_b}{\tilde{v}_{b-n}}$$

denotes the volume of the Stiefel manifold of dimension $n \times b$*, with*

$$\tilde{v}_n = \frac{2^n \pi^{n^2}}{\Gamma_n(n)}$$

being the volume of the unitary group of size n*, and*

$$\widetilde{\mathbf{\Lambda}} = \mathbf{I}_n + \frac{\mathbf{D}^{-1}}{\gamma}.$$

Moreover, $\widetilde{\mathbf{F}}$ *is a square matrix of size* b*, whose generic entry is given by* $\widetilde{\mathbf{F}}_{i,j} = e^{\gamma \widetilde{\lambda}_i y_j^2}$ *if* $1 \leq i \leq n$ *and* $1 \leq j \leq b$*, otherwise* $\widetilde{\mathbf{F}}_{i,j} = (\gamma y_j^2)^{(b-i)}$*, where* y_j^2 *is the* j*-th eigenvalue of* $\mathbf{Y}^\dagger \mathbf{Y}$ *and* $\widetilde{\lambda}_i$ *the* i*-th entry of the diagonal matrix* $\widetilde{\mathbf{\Lambda}}$*.*

Proof. By our assumptions, $\mathbf{Y}|\mathbf{X} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_b + \gamma \mathbf{X}^{\mathsf{H}} \mathbf{X})$, i.e.,

$$p(\mathbf{Y}|\mathbf{X}) = \frac{e^{-\mathrm{Tr}\left(\mathbf{Y}\mathbf{C}^{-1}\mathbf{Y}^{\mathsf{H}}\right)}}{\pi^{mb}|\mathbf{C}|^m}, \tag{4.2}$$

with $\mathbf{C} = \mathbf{I}_b + \gamma \mathbf{X}^{\mathsf{H}} \mathbf{X}$. Replacing $\mathbf{X} = \mathbf{D}^{1/2} \mathbf{\Phi}$ in (4.2) and integrating over $\mathbf{\Phi}$ by exploiting [2, Eq.(57)], the result immediately follows. Notice that the algebra in our derivation is the same as that in [2, Appendix A], even if therein the authors did not refer to channel estimation. The fact that the algebra in the two cases coincides is due to the assumption of having the error matrix variance equal to σ_ϵ^2 and the MMSE estimate variance equal to $1 - \sigma_\epsilon^2$ (see e.g. [54, and references therein]).

Proposition 9. *Given a channel as in (4.1), assume that an instantaneous channel estimate with Gaussian distributed error* $\mathbf{E}$ *is provided at the receiver, i.e.,* $\widehat{\mathbf{H}}$ *is a deterministic matrix. The pdf of the channel output, conditionally on the input matrix* $\mathbf{X}$*, can be expressed as a function of* $\widehat{\mathbf{H}}$*. I.e.,*

$$p(\mathbf{Y}|\mathbf{X}) = e^{\mathrm{Tr}\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}\mathbf{\Lambda}^{-1}\mathbf{\Phi}\right)} \frac{e^{-||\mathbf{Y}||^2} e^{-\mathrm{Tr}\left(\widehat{\mathbf{H}}^{\mathsf{H}}\widehat{\mathbf{H}}\mathbf{D}(\mathbf{I}_n + \gamma\sigma_\epsilon^2\mathbf{D})^{-1}\right)}}{\pi^{mb}|\mathbf{I}_n + \gamma\sigma_\epsilon^2\mathbf{D}|^m} \cdot$$
$$e^{\sqrt{\gamma}\mathrm{Tr}\left(\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}(\mathbf{I}_n - \mathbf{\Lambda}^{-1})\mathbf{D}^{1/2}\widehat{\mathbf{H}}^{\mathsf{H}} + \widehat{\mathbf{H}}\mathbf{D}^{1/2}(\mathbf{I}_n - \mathbf{\Lambda}^{-1})\mathbf{\Phi}\mathbf{Y}^{\mathsf{H}}\right)} \tag{4.3}$$

where

$$\mathbf{\Lambda} = \mathbf{I}_n + \frac{\mathbf{D}^{-1}}{\gamma\sigma_\epsilon^2}.$$

Notice that (4.3) unfortunately does not depend solely on $\mathbf{D}$, since the dependence on the input beamforming matrix $\mathbf{\Phi}$ cannot be averaged out. Indeed, a closed-form evaluation of

$$\int_{\mathcal{S}(n,b)} e^{\sqrt{\gamma}\mathrm{Tr}\left(\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}(\mathbf{I_n}-\mathbf{\Lambda^{-1}})\mathbf{D^{1/2}}\widehat{\mathbf{H}}^{\mathsf{H}}+\widehat{\mathbf{H}}\mathbf{D^{1/2}}(\mathbf{I_n}-\mathbf{\Lambda^{-1}})\mathbf{\Phi}\mathbf{Y}^{\mathsf{H}}\right)} \cdot e^{\mathrm{Tr}\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}\mathbf{\Lambda^{-1}}\mathbf{\Phi}\right)}\,\mathrm{d}\mathbf{\Phi} \tag{4.4}$$

would require the extension to unitary matrices of the result [52, Integral B.1], currently proven only for arbitrary complex matrices. A situation akin to this last one, but even more involved, is met when one eventually assumes that only the channel singular values are reliably estimated, while the left and right singular vectors are not. That is, the average singular values $\widehat{\mathbf{\Sigma}}$ of the channel are given, while the channel singular vectors are assumed to be isotropically distributed on the unitary groups of dimensions given by the number of transmit and receive antennas for $\mathbf{V}$ and $\mathbf{U}$ respectively. In this case, too, in order to perform the average with respect to $\mathbf{V}$, one would need to generalize the result mentioned above[52, Integral B.1] to unitary matrices.

Proof. The result follows by observing that, for deterministic $\widehat{\mathbf{H}}$, $\mathbf{Y}|\mathbf{X}$ is noncentral Gaussian, with correlated entries, i.e.,

$$p(\mathbf{Y}|\mathbf{X}) = \frac{e^{-\mathrm{Tr}\left(\left(\mathbf{Y}-\sqrt{\gamma}\widehat{\mathbf{H}}\,\mathbf{X}\right)\mathbf{C}^{-1}\left(\mathbf{Y}-\sqrt{\gamma}\widehat{\mathbf{H}}\,\mathbf{X}\right)^{\mathsf{H}}\right)}}{\pi^{mb}|\mathbf{C}|^{m}}, \tag{4.5}$$

with $\mathbf{C} = \mathbf{I}_b + \gamma\sigma_\epsilon^2\mathbf{X}^{\mathsf{H}}\mathbf{X}$. Since, by assumption, $n \leq b$,

$$\mathbf{C} = \mathbf{I}_b + \gamma\sigma_\epsilon^2\mathbf{\Phi}^{\mathsf{H}}\mathbf{D}\mathbf{\Phi},$$

and, by matrix inversion Lemma,

$$\mathbf{C}^{-1} = \mathbf{I}_b - \mathbf{\Phi}^{\mathsf{H}}\left(\mathbf{I}_n + \frac{\mathbf{D}^{-1}}{\gamma\sigma_\epsilon^2}\right)^{-1}\mathbf{\Phi},$$

which, replaced in (4.5), together with $\mathbf{X} = \mathbf{D}^{1/2}\mathbf{\Phi}$, leads to (4.3).

Between the two limiting cases of Gaussian and deterministic distributed channel estimate, we next consider the case where $\widehat{\mathbf{H}}$ is a complex random matrix with a smooth distribution. In this case, we are able to give an approximation of the output pdf, up to a scaling factor $c_{\sigma_\epsilon^2}$, depending on the estimation error variance σ_ϵ^2.

Theorem 4. *Given a channel as in (4.1), assume that* $\widehat{\mathbf{H}}$ *is a complex random matrix with a smooth distribution. The pdf of the matrix-variate channel output, conditionally on the input power allocation matrix* $\mathbf{D}$*, can be expressed then as*

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\gamma^{-mn} c_{\sigma_\epsilon^2} |\mathbf{F}| e^{-\|\mathbf{Y}\|^2} \prod_{i=b-n+1}^{b} \Gamma(i) v_{\mathcal{S}}}{\mathcal{V}(\mathbf{Y}^{\mathsf{H}}\mathbf{Y})\mathcal{V}(\mathbf{\Theta})|\mathbf{D}|^m |(\pi\mathbf{\Theta}|^{b-n})} \,. \tag{4.6}$$

In (4.6), $\mathbf{F}$ *is a square matrix of size* b*, whose generic entry* $\mathbf{F}_{i,j} = e^{\theta_i y_j^2}$ *if* $1 \leq i \leq n$ *and* $1 \leq j \leq b$*, otherwise* $\mathbf{F}_{i,j} = (y_j^2)^{(b-i)}$*, with* y_j^2 *the* j*-th eigenvalue of* $\mathbf{Y}^\dagger\mathbf{Y}$*, and* θ_i *the* i*-th element of the diagonal matrix* $\mathbf{\Theta} = \mathbf{\Lambda}^{-1} + \mathbf{I}_n + \gamma\sigma_\epsilon^2\mathbf{D}$ *of size* n*.*

Proof. Following the same arguments as for the previous Proposition, we start by writing

$$p(\mathbf{Y}|\mathbf{X},\widehat{\mathbf{H}}) = \frac{e^{-\mathrm{Tr}\left((\mathbf{Y}-\sqrt{\gamma}\widehat{\mathbf{H}}\,\mathbf{X})\mathbf{C}^{-1}(\mathbf{Y}-\sqrt{\gamma}\widehat{\mathbf{H}}\,\mathbf{X})^{\mathsf{H}}\right)}}{\pi^{mb}|\mathbf{C}|^m} \,, \tag{4.7}$$

again with $\mathbf{C} = \mathbf{I}_b + \gamma\sigma_\epsilon^2\mathbf{X}^{\mathsf{H}}\mathbf{X}$. Let us now define the matrix

$$\mathbf{\Delta} = \mathbf{D}(\mathbf{I}_n + \gamma\sigma_\epsilon^2\mathbf{D})^{-1} \,, \tag{4.8}$$

which will be useful in order to write more compact expressions during the proof. We first expand the product in (4.7), obtaining

$$\begin{aligned} p(\mathbf{Y}|\mathbf{X},\widehat{\mathbf{H}}) = \frac{e^{-||\mathbf{Y}||^2}}{\pi^{mb}|\mathbf{C}|^m} e^{-\gamma\mathrm{Tr}\left(\widehat{\mathbf{H}}^{\mathsf{H}}\widehat{\mathbf{H}}\mathbf{X}\mathbf{C}^{-1}\mathbf{X}^{\mathsf{H}}\right)} \\ e^{\mathrm{Tr}\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}\mathbf{\Lambda}^{-1}\mathbf{\Phi}\right)} e^{\sqrt{\gamma}\mathrm{Tr}\left(\mathbf{Y}\mathbf{C}^{-1}\mathbf{X}^{\mathsf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}+\widehat{\mathbf{H}}\mathbf{X}\mathbf{C}^{-1}\mathbf{Y}^{\mathsf{H}}\right)} \,. \end{aligned} \tag{4.9}$$

Recalling the definition of $\mathbf{C}$,

$$\mathbf{X}\mathbf{C}^{-1}\mathbf{X}^{\mathsf{H}} = \mathbf{\Delta}.$$

To obtain the expression for $p(\mathbf{Y}|\mathbf{D})$, we have to integrate out the dependence on both $\widehat{\mathbf{H}}$, and $\mathbf{\Phi}$.

Indeed, one can average out the dependence on the complex matrix $\widehat{\mathbf{H}}$ by virtue of [52, Integral B.1], thus obtaining

$$\begin{aligned} c_{\sigma_\epsilon^2} \int_{\mathbb{C}} e^{-\gamma\mathrm{Tr}\left(\mathbf{\Delta}\widehat{\mathbf{H}}^{\mathsf{H}}\widehat{\mathbf{H}}\right)} e^{\sqrt{\gamma}\mathrm{Tr}\left(\mathbf{Y}\mathbf{C}^{-1}\mathbf{X}^{\mathsf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}+\widehat{\mathbf{H}}\mathbf{X}\mathbf{C}^{-1}\mathbf{Y}^{\mathsf{H}}\right)} \mathrm{d}\widehat{\mathbf{H}} = \\ c_{\sigma_\epsilon^2}|\mathbf{D}^{-1}\mathbf{C}|^m \left(\frac{\pi}{\gamma}\right)^{mn} e^{\mathrm{Tr}\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}(\mathbf{I}_n+\gamma\sigma_\epsilon^2\mathbf{D})\mathbf{\Phi}\right)} \end{aligned} \tag{4.10}$$

where the input structure has been exploited. From (4.10), we can write that

$$p(\mathbf{Y}|\mathbf{\Phi},\mathbf{D}) = \left(\frac{\pi}{\gamma}\right)^{mn} \frac{e^{-||\mathbf{Y}||^2}}{\pi^{mb}|\mathbf{D}|^m} e^{\mathrm{Tr}\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}(\mathbf{\Lambda}^{-1}+(\mathbf{I}_n+\gamma\sigma_\epsilon^2\mathbf{D}))\mathbf{\Phi}\right)}. \tag{4.11}$$

The next integral is to be performed over the rectangular matrix $\mathbf{\Phi}$. Since only the exponential term depends on $\mathbf{\Phi}$, we can write the integral as

$$\int_{\mathcal{S}(n,b)} e^{\mathrm{Tr}\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}(\mathbf{\Lambda^{-1}}+(\gamma\sigma_\epsilon^{\mathbf{2}})^{\mathbf{2}}\mathbf{D^2}(\mathbf{I_n}+\gamma\sigma_\epsilon^{\mathbf{2}}\mathbf{D})^{\mathbf{-1}})\mathbf{\Phi}\right)} \mathrm{d}\mathbf{\Phi} = \frac{v_{\mathcal{S}} \prod_{i=b-n+1}^{b} \Gamma(i)}{|\mathbf{\Theta}|^{b-n}\mathcal{V}(\mathbf{Y}^{\mathsf{H}}\mathbf{Y})\mathcal{V}(\mathbf{\Theta})}|\mathbf{F}| \tag{4.12}$$

where, again, the result has been obtained following the same approach as in [2, Formula (54)]. Replacing the integration result into (4.11), we finally get (4.6).

We remark that, by following the above steps, when $\widehat{\mathbf{H}}$ is assumed to be Gaussian distributed, we recover the same result as in Proposition 8.

In the case $n = b$, the derivation simplifies, since no limiting procedure has to be performed in order to evaluate the integral over $\mathbf{\Phi}$. Indeed, we can directly apply [29, Formula (92)] to (4.12) thus obtaining

$$\int_{\mathcal{S}(n,b)} e^{\mathrm{Tr}\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}(\mathbf{\Lambda}^{-1}+(\mathbf{I}_n+\gamma\sigma_\epsilon^2\mathbf{D}))\mathbf{\Phi}\right)} \mathrm{d}\mathbf{\Phi} = {}_0F_0\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y},\mathbf{\Theta}\right).$$

Gathering all together, the alternative expression of (4.6) in the case of $n = b$ is given by

$$p(\mathbf{Y}|\mathbf{D}) = \left(\frac{\pi}{\gamma}\right)^{mn} \frac{e^{-||\mathbf{Y}||^2}}{\pi^{mb}|\mathbf{D}|^m} {}_0F_0\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y},\mathbf{\Theta}\right). \tag{4.13}$$

4.4 Conclusion

We provided expressions for the output statistics of block-fading MIMO channels, under the assumption that the receiver is given an channel estimate that is affected by Gaussian error. We derived the pdf of the received signal in several cases of practical interests, including the MMSE estimation, and we presented some open issues regarding the channel output characterization.

Chapter 5

Ergodic Capacity Analysis of MIMO Relay Network over Rayleigh-Rician Channels

In this chapter, we present an analytical characterization of the ergodic capacity for an AF MIMO relay network over asymmetric channels. Indeed, such two-hop system, the source-relay and relay-destination channels undergo Rayleigh and Rician fading, respectively. Considering arbitrary-rank means for the relay-destination channel, we first investigate the marginal distribution of an unordered eigenvalue of the cascaded AF channel, and we provide the analytical expression of the ergodic capacity of the system. The closed-form expressions that we derive are computationally efficient and validated by numerical simulation. Our results also show the impact of the signal-to-noise ratio and of the Rician factor on such asymmetric relay network.

5.1 Introduction

Data transmission through relay channel has been proved to improve coverage, reliability and quality-of-service in wireless systems. Among several proposed relay schemes, amplify-and-forward (AF) has attracted significant attention since it can be easily analyzed and implemented. An AF two-hop system is a classic half-duplex model, where the source sends signal to the relay in the first hop, and then the relay broadcasts the received signal to the destination after a simple amplification. Such a model can then be enhanced by introducing MIMO technology, which can bring remarkable improvements in network performance over conventional single-input and single-output systems.

The performance of AF MIMO relay networks have been widely analyzed by

applying either asymptotic analysis [55, 56] (i.e., assuming an infinite number of antennas or nodes), or finite random matrix theory [57]. What these two approaches have in common is the assumption that both channels in the relay system are subject to Rayleigh fading. In real-world environments, however, the relay node may be deployed closer either to the source or to the destination. In this case, a strong LoS path between the two close-by nodes may exist and the channel on the second hop is affected by Rician, rather than Rayleigh, fading. Note that, such an asymmetric channel model can be seen as a generalization of the traditional two-hop Rayleigh fading channel, since Rayleigh fading can be considered as a limiting case of Rician fading.

In the first of the above asymmetric scenarios (i.e., a relay close to the source), the analysis of the system performance is trivial. Much more challenging, instead, is the case where the channels on the two hops are modeled as Rayleigh and Rician, respectively. This second scenario has indeed attracted significant attention in the literature. In particular, [58] provides the exact expression of the moment generation function and the moments of the instantaneous SNR, under the assumption that an orthogonal space-time blocking coding scheme is applied. In [59], the network performance is studied considering the relay to be equipped with a single antenna. To the best of our knowledge, no analytical expression instead exists for the ergodic capacity in a AF MIMO relay network over asymmetric fading channels in presence of a multiple-antenna relay node. In this letter, we therefore fill this gap. By using finite-dimensional random matrix theory, we provide the closed-form expression of the unordered eigenvalue distribution of the cascaded asymmetric relay channel when the two hops are characterized by Rayleigh and Rician fading, respectively. Through this expression, we also derive the analytical expression of the ergodic capacity of the system, with arbitrary-rank means of Rician channel. Furthermore, by numerical simulation, we investigate the network performance as the Rician factor and the SNR vary.

5.2 System Model

We consider a two-hop relay network where the source, the relay and the destination nodes are equipped with n, r and m antennas, respectively. All nodes operate in half-duplex mode. We assume that no direct link exists between source and destination. The destination has perfect CSI on the source-relay and relay-destination channels, while the source and relay have no CSI.

Following [56, 57], we consider that data transmission takes place in two phases, according to the following scheme. In the first phase, the source transmits signal $\mathbf{x}$, which is a vector with n components, towards the relay. The entries of $\mathbf{x}$ are assumed to be i.i.d., zero-mean, circular symmetric, complex Gaussian random variables and

the power irradiated by each antenna is assumed to be equal to ρ/n, i.e., $\mathbb{E}[\mathbf{x}\mathbf{x}^{\mathsf{H}}] = \rho/n\mathbf{I}_n$, where ρ is the signal-to-noise ratio. In the second phase, the relay simply forwards a scaled version of the signal it received from the source. Let $\mathbf{H}_1 \in \mathbb{C}^{r\times n}$ be the channel matrix between source and relay, and $\mathbf{H}_2 \in \mathbb{C}^{m\times r}$ be the channel matrix between relay and destination. Then, the signal received at the destination can be expressed as

$$\mathbf{y} \;=\; \mathbf{H}_2\mathbf{A}\mathbf{H}_1\mathbf{x} + \mathbf{H}_2\mathbf{A}\mathbf{n}_r + \mathbf{n}_d \tag{5.1}$$

where $\mathbf{A} = \sqrt{a}\mathbf{I}_r$ is an $r \times r$ linear transformation matrix representing the power amplification at the relay, and $\mathbf{n}_r$ and $\mathbf{n}_d$ are, respectively, the noise vectors at the relay and at the destination, whose entries are modeled as i.i.d. zero-mean, unit-variance, Gaussian random variables.

The source-relay channel is assumed to be affected by Rayleigh fading. Thus, the entries of $\mathbf{H}_1$ are i.i.d complex Gaussian random variables with zero mean and unit variance. On the other hand, the relay-destination channel is assumed to be affected by Rician fading so that the entries of $\mathbf{H}_2$ can be written as

$$\mathbf{H}_2 = \sqrt{\frac{\kappa}{\kappa+1}}\bar{\mathbf{H}}_2 + \sqrt{\frac{1}{\kappa+1}}\widetilde{\mathbf{H}}_2 \tag{5.2}$$

where κ is the Rician factor, $\bar{\mathbf{H}}_2$ is deterministic and the entries of $\widetilde{\mathbf{H}}_2$ are i.i.d. complex Gaussian with zero mean and unit variance. For simplicity of notation, we define $\tilde{\kappa} = 1 + \kappa$.

Let $q = \min(r,m)$, $s = \min(n,q)$, $\mathbf{\Lambda} = \mathrm{diag}(\lambda_{\mathbf{1}}, \ldots, \lambda_{\mathbf{q}})$ be the non-zero eigenvalues of $\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}}$, and $\mathbf{H}$ be an $n \times q$ random matrix with i.i.d, circularly symmetric, complex Gaussian entries with zero mean and unit variance. Then, the ergodic capacity of the AF-MIMO relay channel described above is given by [57, eq. (13)]

$$C(\rho) = \frac{s}{2}\int_0^\infty \log_2\left(1 + \frac{\rho a}{n}z\right) p(z)\ \mathrm{d}z \tag{5.3}$$

where z denotes an unordered eigenvalue of the random matrix $\mathbf{Z} = \mathbf{H}^{\mathsf{H}}\mathbf{B}\mathbf{H}$, $p(z)$ denotes the probability density function (pdf) of z, and $\mathbf{B} = \mathbf{\Lambda}(\mathbf{I} + \mathbf{a}\mathbf{\Lambda})^{-\mathbf{1}}$ is a $q \times q$ diagonal matrix.

5.3 Performance Analysis

The ergodic capacity of the above described AF-MIMO channel can be obtained by deriving the closed-form expression of $p(z)$ and plugging it into (5.3). In order to do so, we first assume the LoS component of the Rician channel (i.e., $\bar{\mathbf{H}}_2$) to be full-rank; this case is referred to as "non-i.i.d. Rician fading" in [60] and allows for a relatively simpler analysis (Sec. 5.3.1). Then, we will deal with the case where $\bar{\mathbf{H}}_2$ is low-rank (Sec. 5.3.2).

5.3.1 Closed-form Expression of The Ergodic Capacity

The pdf of the unordered eigenvalue of $\mathbf{Z} = \mathbf{H}^{\mathsf{H}}\mathbf{B}\mathbf{H}$ can be written as

$$p(z) = \int_{\mathbf{B}} p_{z|\mathbf{B}}(z|\mathbf{B}) p_{\mathbf{B}}(\mathbf{B})\, \mathrm{d}\mathbf{B} \tag{5.4}$$

where $p_{z|\mathbf{B}}(z|\mathbf{B})$ is the pdf of z conditioned on $\mathbf{B}$ and can be written as [57, eq. (95)]:

$$p_{z|\mathbf{B}}(z|\mathbf{B}) = \frac{1}{s\mathcal{V}(\mathbf{B})} \sum_{k=q-s+1}^{q} \frac{z^{n+k-q-1}}{\Gamma(n-q+k)} |\mathbf{V}_k| \,. \tag{5.5}$$

In (5.5), $\mathbf{V}$ is a $q \times q$ matrix with entries given by:

$$(\mathbf{V}_k)_{i,j} = \begin{cases} b_i^{q-j}, & q-j+1 \neq k \\ \mathrm{e}^{-z/b_i} b_i^{q-n-1}, & q-j+1 = k \,. \end{cases}$$

The expression of $p_{\mathbf{B}}(\mathbf{B})$ is instead given by the following proposition.

Proposition 10. *Consider a communication system and matrix* $\mathbf{B}$ *as described above. Then, the pdf of* $\mathbf{B}$ *can be written as:*

$$p_{\mathbf{B}}(\mathbf{B}) = \frac{\tilde{\kappa}^{pq} \mathcal{V}(\mathbf{B}) |\mathbf{B}|^{p-q} |\mathbf{I} - a\mathbf{B}|^{-p-1} |\mathbf{F}|}{(p-q)!^q \mathrm{e}^{\mathsf{Tr}\{\tilde{\kappa}\mathbf{B}(\mathbf{I}-a\mathbf{B})^{-1} + \kappa\mathbf{M}\}} \mathcal{V}(\tilde{\kappa}\kappa\mathbf{M})} \tag{5.6}$$

where $\mathbf{F} = \{{}_0F_1(;p-q+1;\kappa\tilde{\kappa}\mu_i b_j/(1-ab_j))\}$, *and* $\mathbf{M} = \bar{\mathbf{H}}_2 \bar{\mathbf{H}}_2{}^{\mathsf{H}}$.

Proof. We first observe that $\mathbf{H}_2 = \widehat{\mathbf{H}}/\sqrt{\tilde{\kappa}}$ where $\widehat{\mathbf{H}} = \sqrt{\kappa}\bar{\mathbf{H}}_2 + \widetilde{\mathbf{H}}_2$ is a standard non-central Wishart matrix with mean $\sqrt{\kappa}\bar{\mathbf{H}}_2$. The joint pdf of the ordered eigenvalues, $\widehat{\mathbf{\Lambda}}$, of $\widehat{\mathbf{H}}^{\mathsf{H}}\widehat{\mathbf{H}}$ is given by [29, eq. (102)]; by using (2.18), it can be written as:

$$p_{\widehat{\mathbf{\Lambda}}}(\widehat{\mathbf{\Lambda}}) = \frac{|\widehat{\mathbf{\Lambda}}|^{p-q} \mathcal{V}(\widehat{\mathbf{\Lambda}}) |\{{}_0F_1(;p-q+1;\kappa\mu_i\hat{\lambda}_j)\}|)}{(p-q)!^q \mathrm{e}^{\mathsf{Tr}\{\widehat{\mathbf{\Lambda}}+\kappa\mathbf{M}\}} \mathcal{V}(\kappa\mathbf{M})} \tag{5.7}$$

where μ_i, $i = 1, \ldots, q$, are the eigenvalues of $\mathbf{M} = \bar{\mathbf{H}}_2 \bar{\mathbf{H}}_2{}^{\mathsf{H}}$ and $\widehat{\mathbf{\Lambda}}$ is a diagonal matrix whose elements are $\hat{\lambda}_j$, $j = 1, \ldots, q$. The pdf of $\mathbf{\Lambda}$ can be then obtained as

$$\begin{aligned} p_{\mathbf{\Lambda}}(\mathbf{\Lambda}) &= \tilde{\kappa}^q p_{\widehat{\mathbf{\Lambda}}}(\tilde{\kappa}\mathbf{\Lambda}) \\ &= \frac{\tilde{\kappa}^{pq} \mathcal{V}(\mathbf{\Lambda}) |\{\mathbf{{}_0F_1}(;\mathbf{p-q+1};\kappa\tilde{\kappa}\mu_\mathbf{i}\lambda_\mathbf{j})\}|}{(p-q)!^q |\mathbf{\Lambda}|^{\mathbf{q-p}} \mathrm{e}^{\mathsf{Tr}\{\tilde{\kappa}\mathbf{\Lambda}+\kappa\mathbf{M}\}} \mathcal{V}(\tilde{\kappa}\kappa\mathbf{M})} \,. \end{aligned} \tag{5.8}$$

Since $\mathbf{B} = \mathbf{\Lambda}(\mathbf{I} + \mathbf{a}\mathbf{\Lambda})^{-\mathbf{1}}$, the pdf of $\mathbf{B}$ can be written as a function of the pdf of $\mathbf{\Lambda}$, i.e., $p_{\mathbf{B}}(\mathbf{B}) = |\mathbf{I} - a\mathbf{B}|^{-2} p_{\mathbf{\Lambda}}(\mathbf{B}(\mathbf{I} - a\mathbf{B})^{-1})$ which yields (5.6). Note that in the derivation of (5.6) we exploited the property $\mathcal{V}(\mathbf{B}(\mathbf{I}-a\mathbf{B})^{-1}) = \mathcal{V}(\mathbf{B})|\mathbf{I}-a\mathbf{B}|^{1-q}$. □

Then, we replace $p_{\mathbf{B}}(\mathbf{B})$ in (5.4) with the expression in (5.6) and obtain the result below.

Proposition 11. *The pdf of an unordered eigenvalue z of $\mathbf{Z} = \mathbf{H}^{\mathsf{H}}\mathbf{B}\mathbf{H}$ is given by*

$$p(z) = \frac{A}{\mathcal{V}(\mathbf{M})} \sum_{k=q-s+1}^{q} \frac{z^{c_k-1}}{\Gamma(c_k)} |\mathbf{W}_k| \tag{5.9}$$

where $c_k = n - q + k$, the constant A is given by

$$A = \frac{\tilde{\kappa}^{pq} \mathrm{e}^{-\kappa \mathsf{Tr}\{\mathbf{M}\}}}{s(p-q)!^q (\kappa\tilde{\kappa})^{q(q-1)/2}}, \tag{5.10}$$

and $\mathbf{W}_k$ is a $q \times q$ matrix whose entries are as follows:

$$(\mathbf{W}_k)_{i,j} = \sum_{h=0}^{j-1} \frac{\binom{j-1}{h} \Gamma(d_{j,h}) {}_1F_1(d_{j,h}; p-q+1; \kappa\mu_i)}{a^{-h}\tilde{\kappa}^{d_{j,h}}} \tag{5.11}$$

for $i,j = 1,\ldots,q$, $j \neq q-k+1$ and

$$(\mathbf{W}_k)_{i,j} = \frac{2\mathrm{e}^{-za}}{\tilde{\kappa}^{p-n}} \sum_{\ell=0}^{\infty} \sum_{h=0}^{n} \frac{\binom{n}{h} (\kappa\mu_i)^{\ell} (z\tilde{\kappa})^{g_{\ell,h}/2}}{a^{-h}\tilde{\kappa}^h \ell! (p-q+1)_{\ell}} K_{g_{\ell,h}}(2\sqrt{z\tilde{\kappa}}) \tag{5.12}$$

for $i = 1,\ldots,q$, $j = q-k+1$, where $d_{j,h} = p+1-j+h$ and $g_{\ell,h} = p-n+\ell+h$. In (5.12), $K_v(x)$ denotes the modified Bessel function of the second kind.

Proof. As mentioned above, $p(z)$ can be computed by using (5.6) in (5.4). As for the integration domain, we observe that the i-th eigenvalue of $\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}}$, λ_i, is such that $0 \leq \lambda_i < +\infty$. Thus, $b_i = \lambda_i/(1+a\lambda_i)$ has support in [0,1/a]. Moreover, the expression of $p_{\mathbf{B}}(\mathbf{B})$ provided in (5.6) refers to the *ordered* eigenvalue distribution of $\mathbf{B}$, hence the integral in (5.4) should be taken under the constraint $0 \leq b_q < \ldots < b_1 \leq 1/a$. By substituting (5.5) and (5.6) in (5.4), we obtain:

$$p(z) = \sum_{k=q-s+1}^{q} \frac{A z^{c_k-1}}{\mathcal{V}(\mathbf{M})\Gamma(c_k)} \int \frac{|\mathbf{I}-a\mathbf{B}|^{-p-1}|\mathbf{F}||\mathbf{V}_k|\,\mathrm{d}\mathbf{B}}{|\mathbf{B}|^{q-p}\mathrm{e}^{\mathsf{Tr}\{\tilde{\kappa}\mathbf{B}(\mathbf{I}-a\mathbf{B})^{-1}\}}} \tag{5.13}$$

$$= \frac{A}{\mathcal{V}(\mathbf{M})} \sum_{k=q-s+1}^{q} \frac{z^{c_k-1}}{\Gamma(c_k)} |\mathbf{W}_k| \tag{5.14}$$

where the constant A is given by (5.10) and $c_k = n-q+k$. The $q \times q$ matrix $\mathbf{W}_k$ appearing in (5.14) derives from the application of [17, Corollary 2] to the integral

in (5.13) and its entries are given by

$$(\mathbf{W}_k)_{i,j} = \begin{cases} \displaystyle\int_0^{\frac{1}{a}} \frac{\mathrm{e}^{\frac{-\tilde{\kappa}x}{1-ax}} x^{p-j} f_{ij}}{(1-ax)^{p+1}}\,\mathrm{d}x, & j \neq q-k+1 \\ \displaystyle\int_0^{\frac{1}{a}} \frac{\mathrm{e}^{\frac{-\tilde{\kappa}x}{1-ax}-z/x} f_{ij}}{x^{1-p+n}(1-ax)^{p+1}}\,\mathrm{d}x, & j = q-k+1 \end{cases} \tag{5.15}$$

where $f_{ij} = {}_0F_1(;p-q+1;\kappa\tilde{\kappa}\mu_i x/(1-ax))$ has been defined below (5.6). The integrals in (5.15) can be solved by a suitable change of the integration variable and by expanding the powers of the binomial $\tilde{\kappa}+ax$. As a result, after some computations (omitted for lack of space), the matrix $\mathbf{W}_k$ can be rewritten as in (5.11) and (5.12). □

Eventually, the analytical expression of ergodic capacity can be obtained by substituting (5.9) into (5.3).

5.3.2 Low Rank LoS Rician Fading Component

We now consider the case where the LoS component of the Rician channel, $\bar{\mathbf{H}}_2$, does not have full rank, i.e., the terms $|\mathbf{W}_k|$ and $\mathcal{V}(\mathbf{M})$, respectively at the numerator and denominator of (5.9), vanish thus leading to a 0/0 indeterminate form. In order to circumvent this problem, a limit must be taken, which can be evaluated by using l'Hôpital's rule. In particular, in the following we assume $\bar{\mathbf{H}}_2{}^{\mathsf{H}}\bar{\mathbf{H}}_2$ to have $0<g<q$ non-zero eigenvalues, i.e., $\mu_{g+1}=\mu_{g+2}=\cdots=\mu_q=0$. Then, the pdf of an unordered eigenvalue z of $\mathbf{Z}=\mathbf{H}^{\mathsf{H}}\mathbf{B}\mathbf{H}$ can be derived by taking the following limit:

$$\begin{aligned} p(z)_{\text{low}} &= \lim_{\mu_{g+1},\ldots,\mu_q\to 0} p(z) \\ &= A\sum_{k=q-s+1}^{q} \frac{z^{c_k+1}}{\Gamma(c_k)} \lim_{\mu_{g+1},\ldots,\mu_q\to 0} \frac{|\mathbf{W}_k|}{\mathcal{V}(\mathbf{M})} \\ &= \frac{\pi_q \Gamma_g(q)}{\pi_g \Gamma_q(q)} \frac{|\widetilde{\mathbf{W}}_k|}{\mathcal{V}(\widetilde{\mathbf{M}})} |\widetilde{\mathbf{M}}|^{g-q} \end{aligned} \tag{5.16}$$

where we used the result reported in (2.9). In (5.16), $\widetilde{\mathbf{M}}$ is a $g\times g$ matrix with eigenvalues $\mu_1,\ldots,\mu_g$ and $\widetilde{\mathbf{W}}_k$ is a $q\times q$ matrix whose (i,j)-th entry is given by $(\mathbf{W}_k)_{i,j}$, for $i=1,\ldots,g$ and $j=1,\ldots,q$;

$$\left(\widetilde{\mathbf{W}}_k\right)_{i,j} = \sum_{h=0}^{j-1} \frac{\binom{j-1}{h}\Gamma(d_{j,h}+q-i)\kappa^{q-i}}{(p-q+1)_{q-i}a^{-h}\tilde{\kappa}^{d_{j,h}}}$$

for $i = g+1,\ldots,q$, and $j \neq q-k+1$; and

$$\left(\widetilde{\mathbf{W}}_k\right)_{i,j} = \frac{2\mathrm{e}^{-za}}{\tilde{\kappa}^{p-n}} \sum_{h=0}^{n} \frac{\binom{n}{h}\kappa^{q-i}(z\tilde{\kappa})^{g_{q-i,h}/2}}{a^{-h}\tilde{\kappa}^{h}(p-q+1)_{q-i}} K_{g_{q-i,h}}(2\sqrt{z\tilde{\kappa}})$$

for $i = g+1,\ldots,q$, $j = q-k+1$. This result has been obtained by deriving $q-j$ times (with respect to μ_j) the (i,j)-th entry of $\mathbf{W}_k$ and by evaluating it in $\mu_j = 0$, for $i = 1,\ldots,q$ and $j = g+1,\ldots,q$. The other entries of $\mathbf{W}_k$ have been left unchanged. Again, the expression of ergodic capacity with low-rank LoS Rician fading component can be obtained by substituting (5.16) into (5.3).

5.4 Numerical Results

We now validate our analytical derivations through Monte Carlo simulation, and show the information theoretic performance of the channel of (5.1). We consider $n = 2$, $r = 3$, $m = 4$, and the Rician channel with LoS matrix that satisfies $\mathsf{Tr}\{\bar{\mathbf{H}}_2{}^{\mathsf{H}}\bar{\mathbf{H}}_2\} = rm$. More specifically, in Figs. 5.1 and 5.2 we have employed both the full-rank LoS matrix $\bar{\mathbf{H}}_2^f$ and low-rank LoS matrix $\bar{\mathbf{H}}_2^l$, given by

$$\bar{\mathbf{H}}_2^f = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & \sqrt{2} & 2 \\ 1 & 1 & \sqrt{2} \\ 1 & \sqrt{3} & 1 \\ 1 & \sqrt{2} & \sqrt{5} \end{bmatrix}; \bar{\mathbf{H}}_2^l = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Also, while computing the integrals in (5.3) and (5.4), we upper bound the integration variables z and λ so as to obtain an expression that can be efficiently computed. In numerical simulation, an accurate result can be obtained already for $\lambda \leq 20$ and $z \leq 100$.

Figure 5.1 shows the excellent match between the pdf of z, $p(z)$, computed through the analytical expression in (5.4) and the results obtained via Monte Carlo simulation. In particular, considering the full-rank LoS component, Figure 5.1(a) presents the results for different values of the Rician factor κ as well as SNR (ρ), and compares our result to that provided in [57] for the Rayleigh-Rayleigh channel. As expected, as κ decreases, $p(z)$ converges to that obtained for Rayleigh-Rayleigh relay channel ($\kappa = 0$). Figure 5.1(b) presents the same comparison between analytical and simulation results considering low-rank LoS component. Compared with $p(z)$ in full-rank LoS case, Figure 5.1(b) shows a higher concentration of z with smaller value, which incurs the lower ergodic system capacity in low-rank LoS case that can be observed in Figure 5.2.

Figure 5.2 depicts the ergodic capacity of the relay system computed through (5.3) and via Monte Carlo simulation, for different values of the Ricean factor and rank of

LoS component. Again, analytical and numerical results are remarkably close. The fact that the lower the Rician factor is, the higher the ergodic capacity becomes, confirms the validity of our results.

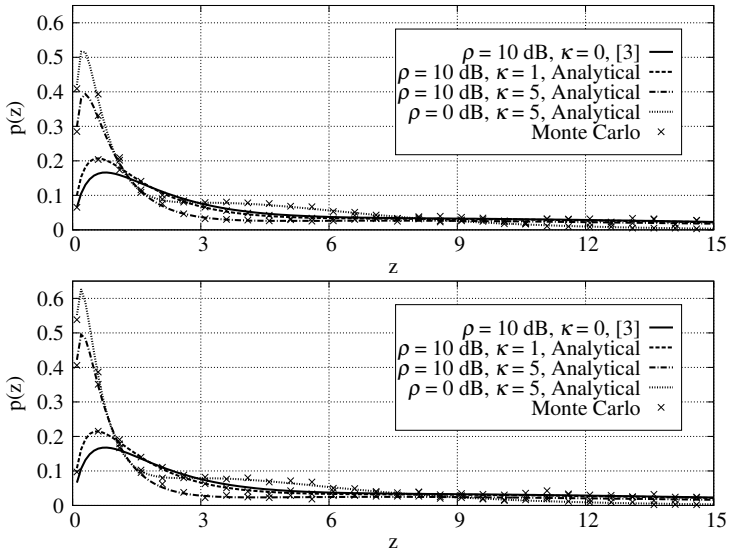

Figure 5.1: Comparison between exact analysis and Monte Carlo simulation: pdf of an unordered eigenvalue z, for full-rank LoS component (top) and low-rank LoS component (bottom).

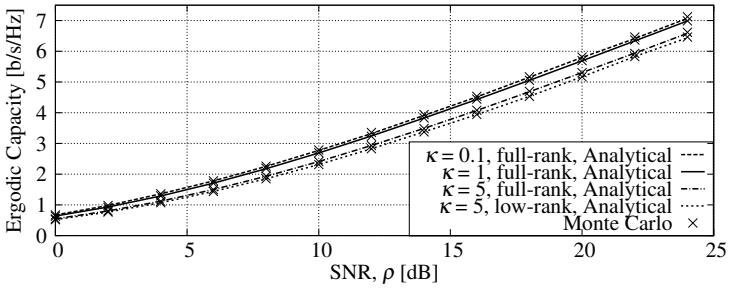

Figure 5.2: Comparison between exact analysis and Monte Carlo simulation: Ergodic capacity vs. SNR with different values of the Rician factor and rank of LoS component.

5.5 Conclusion

We have investigated the ergodic capacity performance of AF MIMO relay networks over asymmetric Rayleigh-Rician channels. In the two cases where the Rician channel has full-rank and low-rank means, we derive the closed-form expression for the marginal pdf of an unordered eigenvalue of the cascaded AF channel. Using these analytical expressions, we derived the ergodic capacity of the system. Our

analysis was validated by showing the excellent matching existing between the results obtained through our exact expressions and those obtained via Monte Carlo simulation.

Chapter 6

Conclusions and Future Work

The topic of this thesis is the application of finite random matrix theory in multiple-antenna wireless communication systems. The contribution can be summarized into two categories: the first is to provide the closed-form output statistics of MIMO block-fading channels with respect to several channel fading models, while the second is to analyze the ergodic capacity of MIMO relay channels over asymmetric fading channels. In the following, we summarize the contributions and discuss future directions of research.

6.1 Summary of Contributions

New results on the finite random matrix theory have been presented in Chapter 2. Prior to the description of the novel expressions, a brief overview of the finite random matrix theory has been given, where fundamental definitions, the matrix-variate distribution and the eigenvalue distribution in random matrix theory have been declared.

Chapter 3 has presented compact expressions of output density in MIMO block-fading channels, paving the way to a more detailed analytical information-theoretic exploration of communications in presence of block-fading. Specifically, the entropy of the unconditional and conditional output have been computed through Monte-Carlo method using the corresponding output pdf. After that, the mutual information of the system has been obtained. The analysis has been performed under the assumptions of two different input structures and absence of any CSI at the transmitter or the receiver.

- When the channel is fed by the i.i.d. Gaussian input, we have focused on the analysis of some classes of channel matrices whose Gramian is unitarily invariant. Although the i.i.d. Gaussian input is not generally capacity achieving input without perfect CSI, the mutual information obtained represents the

highest spectral efficiency that can be obtained using Gaussian codebooks.

  - Considering that the channel impairment is dominated by AWGN noise, we have derived the unconditional output pdf with the channels subject to Rayleigh fading, Rician fading and LMS characterization. Analysis of the output statistics are been performed with arbitrary number of antennas in the system. Note that the output pdf for MIMO block-fading Rayleigh channels has been derived in [35].
  - Considering that the channel impairment is dominated by the co-channel interference, we have evaluated the output pdf when a whitening filter is applied to the received signal. We have considered two fading channel scenarios where the desired signal undergoes Rayleigh channel and Rician channel.

- In the high-SNR regime, the capacity achieving input of the block-fading channels without perfect CSI has been derived as a product between a diagonal matrix and a Stiefel matrix. The diagonal elements are determined by the relationship between the coherence length and the sum of terminal antennas.

  - When the coherence length is larger than the sum of terminal antennas, the diagonal entries are almost surely equal to 1. We have derived the unconditional output pdf as well as the conditional output pdf, separately.
  - When the coherence length is less than the sum of terminal antennas, the diagonal entries are distributed as Beta distribution. The corresponding unconditional output pdf has been derived. The mutual information computed through the closed output statistics are tightly aligned with the approximation expression given in [2].

Chapter 4 has provided the output statistics of MIMO block-fading channels with imperfect CSI. The CSI received by the receiver is impaired by AWGN Gaussian noise and the input is assumed as a product form, which is proven to be capacity achieving in block-fading channels without CSI.

- When the entries of the estimated channel behave as i.i.d. Gaussian distribution, the output pdf conditional on the diagonal matrix in the input has been given.

- When the entries of the estimated channel is assumed to be known with a Gaussian distributed error, the output pdf conditional on the input signal has been given.

- When the estimated channel is subject to a smooth distribution, the output pdf conditional on the diagonal matrix in the input has been given. Note that when the estimated is subject to Gaussian distribution, this general result turns out to be the result given in the first case.

Chapter 5 has presented the ergodic capacity analysis of MIMO relay network over asymmetric fading channels. In this AF relay network, the first channel between the source and the relay is subject to Rayleigh fading, while the second channel between the relay and the destination is subject to Rician fading.

- The unordered eigenvalue distribution of the cascaded asymmetric relay channel has been first characterized, with arbitrary number of antennas in each terminal.
- The ergodic capacity of such relay network over Rayleigh-Rician channels has been derived in closed form. The results have been validated by comparison with the numerical simulation.
- Considering that the LoS component of the Rician channel is not full rank, the unordered eigenvalue distribution of the cascaded relay channel has been given and the corresponding ergodic capacity has been also derived.

6.2 Future Work

Future research directions include the analysis of the two-way relay networks with multiple-antenna implemented in each terminal.

The first extension is to consider the MRT analysis in hop-by-hop MIMO relay networks [61]. When CSI is perfectly known at the transmission and at the receiver, MRT is a scheme that is particularly robust against the severe effects of fading. In MRT, the signal is transmitted along the strongest eigenmode and the received signals are combined using maximal ratio combining. When CSI is only known to the transmitter and the receiver, the outage probability, bit error rate (BER), and the ergodic capacity have been derived in closed form in [62], considering that the dual-hop channels undergo Rayleigh fading and Rician fading. When the relay as well as two terminals are able to obtain the CSI of the two channels, both source and relay could perform beamforming, and the network performance is analyzed in [63]. However, in [63], both dual-hop channels are considered as Rayleigh fading. This assumption intrigues us to explore the network performance of hop-by-hop beamforming relay networks over asymmetric channels. Moreover, similar to [63], we could take into account the effect of imperfect CSI on the network performance.

The second extension is to consider the scenario where multiple destinations exist in relay networks. With multiple users behaving as the destination nodes and CSI

of the second channels is available at relay, opportunistic scheduling is a technique to select a single destination with the highest instantaneous SNR in the second hop channel. In order to maximize the instantaneous SNR, MRT is employed in each terminal. In the literature, [64] exploits the impact of opportunistic scheduling in such networks, considering the antenna selection scheme. In [65], outage probability and average symbol error rate (SER) are analyzed in interference-limited regime, considering single antenna at relay and multiple antennas at source and destinations. The limitation of the existing work is the lack of analysis on the impact of multiple antenna employed at the relay, taking into account opportunistic scheduling, MRT, and imperfect CSI. An interesting research line is therefore the study of the network performance under aforementioned network models, with the help of finite RMT.

Chapter 7

Acronyms

AF	Amplify-and-Forward.
AWGN	Additive White Gaussian Noise.
BER	Bit Error Rate.
b/s/Hz	Bits per Second per Hertz.
cdf	Cumulative Distribution Function.
CSI	Channel State Information.
DF	Decode-and-Forward.
i.i.d.	Independent and Identically Distributed.
LMS	Land Mobile Satellite.
LoS	Line of Sight.
MIMO	Multiple Input and Multiple Output.
MISO	Multiple Input and Single Output.
MMSE	Minimum Mean Squared Error.
MRC	Maximum Ratio Combining.
MRT	Maximum Ratio Transmission.
pdf	Probability Density Function.
RMT	Random Matrix Theory.
SINR	Signal to Interference-Plus-Noise Ratio.
SIMO	Single Input and Multiple Output.
SISO	Single Input and Single Output.
SNR	Signal to Noise Ratio.
STC	Space Time Codes.
SVD	Singular Value Decomposition.

VBLAST	Vertical Bell Laboratories Layered Space Time.
WLAN	Wireless Local Area Network.

Appendix A

Appendix

A.1 Proof of Theorem 1

Let $\mathbf{H}_1$ and $\mathbf{H}_2$ be, respectively, an $m \times n$ and an $m \times p$ ($m \leq p$) Gaussian complex random matrix whose columns are independent, have zero mean, and covariance $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$, respectively.

- For $m \leq n$, the distribution of the ordered eigenvalues of $(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}\mathbf{H}_1\mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}$ is given by [29, eq. (98)]

$$p(\mathbf{\Lambda}) = \frac{\pi_m^2 \mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}|\{{}_1F_0(p+n-m+1;\,;-\frac{\lambda_j}{\omega_i})\}|}{(p+n-m)!^{-m}\Gamma_m(p)\Gamma_m(n)|\mathbf{\Omega}|^n\mathcal{V}(-\mathbf{\Omega}^{-1})} \tag{A.1}$$

  where $\mathbf{\Omega} = \mathbf{\Theta}_1\mathbf{\Theta}_2^{-1}$, and $\omega_1, \ldots, \omega_m$ are the eigenvalues of $\mathbf{\Omega}$. When $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ are scalar matrices, and $\mathbf{\Omega} = \mathbf{\Theta}_1\mathbf{\Theta}_2^{-1} = \omega\mathbf{I}$, the distribution of $\mathbf{\Lambda}$ can be obtained first by applying the limit (2.7) to (A.1):

$$\begin{aligned} p(\mathbf{\Lambda}) &= \frac{\pi_m^2(p+n-m)!^m}{\Gamma_m(p)\Gamma_m(n)}\frac{\mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}}{\omega^{mn}} \\ &\quad \cdot \lim_{\mathbf{\Omega}\to\omega\mathbf{I}} \frac{|\{{}_1F_0(p+n-m+1;\,;-\lambda_j/\omega_i)\}|}{\mathcal{V}(-\mathbf{\Omega}^{-1})} \\ &= \frac{\pi_m^2\Gamma_m(p+n)\mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}}{\Gamma_m(m)\Gamma_m(p)\Gamma_m(n)\omega^{mn}} \\ &\quad \cdot|\{\lambda_j^{m-i}{}_1F_0(p+n-j+1;\,;-\lambda_j/\omega)\}| \end{aligned} \tag{A.2}$$

and then by observing that

$$\begin{aligned}
&|\{\lambda_j^{m-i}{}_1F_0(p+n-i+1;\,;-\lambda_j/\omega)\}| \\
&\quad= |\{\lambda_j^{m-i}(1+\lambda_j/\omega)^{-(p+n-i+1)}\}| \\
&\quad= |\{[\lambda_j/(1+\lambda_j/\omega)]^{m-i}(1+\lambda_j/\omega)^{-(p+n-m+1)}\}| \\
&\quad= \mathcal{V}(\mathbf{\Lambda}(\mathbf{I}+\mathbf{\Lambda}/\omega))|\mathbf{I}+\mathbf{\Lambda}/\omega|^{-(p+n-m+1)} \\
&\quad= \mathcal{V}(\mathbf{\Lambda})|\mathbf{I}+\mathbf{\Lambda}/\omega|^{1-m}|\mathbf{I}+\mathbf{\Lambda}/\omega|^{-(p+n-m+1)} \\
&\quad= \mathcal{V}(\mathbf{\Lambda})|\mathbf{I}+\mathbf{\Lambda}/\omega|^{-(p+n)}\,. \qquad \text{(A.3)}
\end{aligned}$$

- For $m > n$, the distribution of the $n\times n$ random matrix $\mathbf{W} = \mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1}\mathbf{H}_1$ is given by [66, eq. (61)]

$$p(\mathbf{W}) = \frac{\Gamma_m(n+p)\,{}_1F_0(p+n;\,;\mathbf{I}-\mathbf{\Omega}^{-1},\mathbf{W}(\mathbf{I}+\mathbf{W})^{-1})}{\Gamma_m(p)\Gamma_n(m)|\mathbf{\Omega}|^n|\mathbf{W}|^{n-m}|\mathbf{I}+\mathbf{W}|^{p+n}} \qquad \text{(A.4)}$$

where $\mathbf{\Omega} = \mathbf{\Theta_1^{1/2}\Theta_2^{-1}\Theta_1^{1/2}}$ is of size $m \times m$. Note that, for any unitary matrix $\mathbf{V}$ independent of $\mathbf{W}$, we have $|\mathbf{VWV}^{\mathsf{H}}| = |\mathbf{W}|$, $|\mathbf{I}+\mathbf{VWV}^{\mathsf{H}}| = |\mathbf{I}+\mathbf{W}|$. Moreover, the eigenvalues of $\mathbf{VWV}^{\mathsf{H}}(\mathbf{I}+\mathbf{VWV}^{\mathsf{H}})^{-1}$ are the same as those of $\mathbf{W}(\mathbf{I}+\mathbf{W})^{-1}$. Thus, $p(\mathbf{VWV}) = p(\mathbf{W})$. It follows that $\mathbf{W}$ is unitarily invariant.

Let $\mathbf{\Psi} = \mathbf{W(I+W)^{-1}}$ and the eigenvalues of $\mathbf{W}$ be $\lambda_1,\ldots,\lambda_n$. Then, $\mathbf{\Psi}$ has eigenvalues $\psi_j = \lambda_j/(1+\lambda_j)$, for $j = 1,\ldots,n$. In order to compute the hypergeometric function of two matrix arguments of different size appearing in (A.4), we extend $\mathbf{\Psi}$ to the $m \times m$ matrix $\widetilde{\mathbf{\Psi}}$ given by

$$\widetilde{\mathbf{\Psi}} = \begin{bmatrix} \mathbf{\Psi} & \mathbf{0} \\ \mathbf{0} & \mathbf{E} \end{bmatrix}$$

where $\mathbf{E}$ is an $(m-n)\times(m-n)$ matrix whose eigenvalues are $\mathbf{e} = [e_1,\ldots,e_{m-n}]^{\mathsf{T}}$. Then the eigenvalues of $\widetilde{\mathbf{\Psi}}$ are $\tilde{\boldsymbol{\psi}} = [\psi_1,\ldots,\psi_n,e_1,\ldots,e_{m-n}]^{\mathsf{T}}$. It follows that

$$\begin{aligned}
&{}_1F_0(p+n;\,;\mathbf{I}-\mathbf{\Omega^{-1}},\mathbf{\Psi}) \\
&\quad= \lim_{\mathbf{e}\to\mathbf{0}} {}_1F_0(p+n;\,;\mathbf{I}-\mathbf{\Omega^{-1}},\widetilde{\mathbf{\Psi}}) \\
&\quad\overset{(a)}{=} \lim_{\mathbf{e}\to\mathbf{0}} \frac{\Gamma_m(m)(p+n-m)!^m}{\Gamma_m(p+n)} \frac{|\{f_i(\tilde{\psi}_j)\}|}{\mathcal{V}(\mathbf{I}-\mathbf{\Omega}^{-1})\mathcal{V}(\widetilde{\mathbf{\Psi}})} \\
&\quad= \frac{\Gamma_m(m)(p+n-m)!^m}{\Gamma_m(p+n)\mathcal{V}(\mathbf{I}-\mathbf{\Omega}^{-1})} \lim_{\mathbf{e}\to\mathbf{0}} \frac{|\{f_i(\tilde{\psi}_j)\}|}{\mathcal{V}(\widetilde{\mathbf{\Psi}})} \qquad \text{(A.5)}
\end{aligned}$$

where in (a) we applied (2.18) and $f_i(\tilde{\psi}_j) = {}_1F_0(p+n-m+1;;(1-\omega_i^{-1})\tilde{\psi}_j)$, $i,j = 1,\ldots,m$. By applying the limit in (2.7) and the properties in (**??**) and (2.20), the limit in (A.5) can be computed as

$$\lim_{\mathbf{e}\to\mathbf{0}} \frac{|\{f_i(\tilde{\psi}_j)\}|}{\mathcal{V}(\widetilde{\mathbf{\Psi}})} = \frac{\Gamma_m(p+n)\Gamma_n(m)}{\Gamma_n(p+n)\Gamma_m(m)} \frac{(p+n-m)!^{n-m}|\mathbf{F}|}{|\mathbf{\Psi}|^{m-n}\mathcal{V}(\mathbf{\Psi})} \tag{A.6}$$

where for $i = 1,\ldots,m$

$$(\mathbf{F})_{ij} = \begin{cases} f_i(\psi_j), & j = 1,\ldots,n \\ (1-\omega_i^{-1})^{m-j} & j = n+1,\ldots,m \end{cases}$$

We now write the eigenvalue decomposition of $\mathbf{W}$ as $\mathbf{W} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$, where $(\lambda_1,\ldots,\lambda_n) = \mathrm{diag}(\mathbf{\Lambda})$. We then observe that $|\mathbf{W}| = |\mathbf{\Lambda}|$, $|\mathbf{I}+\mathbf{W}| = |\mathbf{I}+\mathbf{\Lambda}|$, $|\mathbf{\Psi}| = |\mathbf{\Lambda}||\mathbf{I}+\mathbf{\Lambda}|^{-\mathbf{1}}$, and that $\mathcal{V}(\mathbf{\Psi}) = \mathcal{V}(\mathbf{\Lambda})|\mathbf{I}+\mathbf{\Lambda}|^{\mathbf{1-n}}$. Therefore, by using (A.4), (A.5), and (A.6), the pdf of $\mathbf{W}$ can be rewritten as

$$p(\mathbf{W}) = \frac{\Gamma_m(p+n)(p+n-m)!^n}{\Gamma_n(p+n)\Gamma_m(p)} \frac{|\mathbf{I}+\mathbf{\Lambda}|^{m-p-n-1}|\mathbf{F}|}{|\mathbf{\Omega}|^n\mathcal{V}(\mathbf{I}-\mathbf{\Omega}^{-1})\mathcal{V}(\mathbf{\Lambda})}\,. \tag{A.7}$$

The pdf of the ordered eigenvalues of a complex random $n \times n$ matrix $\mathbf{W}$ is given by [29, eq. (93)]:

$$p(\mathbf{\Lambda}) = \frac{\pi_n^2\mathcal{V}(\mathbf{\Lambda})^2}{\Gamma_n(n)} \int p_{\mathbf{W}}(\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}})\,\mathrm{d}\mathbf{U}\,.$$

In our case, since $p(\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}})$ does not depend on $\mathbf{U}$, we obtain (2.35).

A.2 Proof of Theorem 2

- For $m \le n$, let us define $\mathbf{W}_2 = \mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}}$. Then the matrix $\mathbf{W} = \mathbf{W}_2^{-1/2}\mathbf{H}_1\mathbf{H}_1{}^{\mathsf{H}}\mathbf{W}_2^{-1/2}$ can be rewritten as $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$ where $\mathbf{H} = \mathbf{W}_2^{-1/2}\mathbf{H}_1$. For any given matrix $\mathbf{W}_2$, for $\mathbf{\Theta} = \theta\mathbf{I}$ and $\mathbf{M}\mathbf{M}^{\mathsf{H}} = \mu\mathbf{I}$, $\mathbf{H}$ is a Gaussian complex matrix with average $\widetilde{\mathbf{M}} = \mathbf{W}_2^{-1/2}\mathbf{M}$ and independent columns whose covariance is $\mathbf{\Sigma} = \theta\mathbf{W}_{\mathbf{2}}^{-\mathbf{1}}$. It follows that, given $\mathbf{W}_2$, $\mathbf{W}$ is a non central Wishart matrix and $p(\mathbf{W}|\mathbf{W}_2)$ is given by [29, eq. (99)]

$$\begin{aligned} p_{\mathbf{W}|\mathbf{W}_2}(\mathbf{W}|\mathbf{W}_2) &= {}_0F_1(;n;\mathbf{\Sigma}^{-\mathbf{1}}\widetilde{\mathbf{M}}\widetilde{\mathbf{M}}^{\mathsf{H}}\mathbf{\Sigma}^{-\mathbf{1}}\mathbf{W}) \\ &\quad \cdot \frac{\mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Sigma}^{-\mathbf{1}}\mathbf{W}\}}|\mathbf{W}|^{n-m}}{\mathrm{e}^{\mathsf{Tr}\{\mathbf{\Sigma}^{-\mathbf{1}}\widetilde{\mathbf{M}}\widetilde{\mathbf{M}}^{\mathsf{H}}\}}\Gamma_m(n)|\mathbf{\Sigma}|^{\mathbf{n}}} \\ &= {}_0F_1(;n;\theta^{-2}\mu\mathbf{W}_2\mathbf{W}) \\ &\quad \cdot \frac{\mathrm{e}^{-\mathsf{Tr}\{\theta^{-1}\mathbf{W}_2\mathbf{W}\}}|\mathbf{W}_2|^n}{\mathrm{e}^{\mu m/\theta}\theta^{nm}\Gamma_m(n)|\mathbf{W}|^{m-n}}\,. \end{aligned} \tag{A.8}$$

On the other hand, $\mathbf{W}_2$ is a central Wishart with covariance $\theta\mathbf{I}$. Thus, the density of $\mathbf{W}$ can be written as

$$\begin{aligned} p_{\mathbf{W}}(\mathbf{W}) &= \int p_{\mathbf{W}|\mathbf{W}_2}(\mathbf{W}|\mathbf{W}_2) p_{\mathbf{W}_2}(\mathbf{W}_2)\,\mathrm{d}\mathbf{W}_2 \\ &= \int {}_0F_1(;n;\frac{\mu}{\theta^2}\mathbf{W}_2\mathbf{W}) \\ &\qquad \cdot \frac{\mathrm{e}^{-\mathsf{Tr}\{\mathbf{W}_2(\mathbf{I}+\mathbf{W})/\theta\}}|\mathbf{W}_2|^{p+n-m}\,\mathrm{d}\mathbf{W}_2}{\mathrm{e}^{\mu m/\theta}\theta^{(p+n)m}\Gamma_m(n)\Gamma_m(p)|\mathbf{W}|^{m-n}} \\ &= \frac{|\mathbf{W}|^{n-m}}{\mathrm{e}^{\mu m/\theta}\theta^{(p+n)m}\Gamma_m(n)\Gamma_m(p)} \\ &\qquad \cdot \int \frac{{}_0F_1(;n;\frac{\mu}{\theta^2}\mathbf{W}_2\mathbf{W})\,\mathrm{d}\mathbf{W}_2}{\mathrm{e}^{\mathsf{Tr}\{\mathbf{W}_2(\mathbf{I}+\mathbf{W})/\theta\}}|\mathbf{W}_2|^{m-p-n}}\,. \end{aligned} \tag{A.9}$$

In order to solve the above integral, we employ the following result

$$\int\limits_{\mathbf{B}=\mathbf{B}^{\mathsf{H}}>0} \frac{{}_pF_q(\mathbf{a};\mathbf{b};\mathbf{C}\mathbf{B})}{|\mathbf{B}|^{m-c}\mathrm{e}^{\mathsf{Tr}\{\mathbf{A}\mathbf{B}\}}}\,\mathrm{d}\mathbf{B} = \frac{{}_{p+1}F_q(\mathbf{a},c;\mathbf{b};\mathbf{C}\mathbf{A}^{-1})}{[\Gamma_m(c)]^{-1}|\mathbf{A}|^c}\,,$$

which holds for $m\times m$ matrices $\mathbf{A}$,$\mathbf{B}$, and $\mathbf{C}$, and for $\mathbb{R}(c) > m-1$ [67, eq. (115)]. Then,

$$p_{\mathbf{W}}(\mathbf{W}) = \frac{\Gamma_m(p+n)\,{}_1F_1(p+n;n;\frac{\mu}{\theta}\mathbf{W}(\mathbf{I}+\mathbf{W})^{-1})}{\Gamma_m(n)\Gamma_m(p)\mathrm{e}^{\mu m/\theta}|\mathbf{W}|^{m-n}|\mathbf{I}+\mathbf{W}|^{p+n}}\,.$$

It can be observed that $p(\mathbf{W})$ depends only on the eigenvalues of $\mathbf{W}$, thus it is unitarily invariant. It follows that the pdf of the ordered eigenvalues of $\mathbf{W}$ is given by [29, eq. (93)]

$$p(\mathbf{\Lambda}) = \frac{\pi_m^2\mathcal{V}(\mathbf{\Lambda})^2}{\Gamma_m(m)}\int p_{\mathbf{W}}(\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}})\,\mathrm{d}\mathbf{U}\,,$$

which provides the results in (2.36).

- For $m > n$, the distribution of the $n\times n$ matrix $\mathbf{W} = \mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1}\mathbf{H}_1$ is given by [29, eq. (105)]

$$p(\mathbf{W}) = \frac{\Gamma_n(p+n)\,{}_1F_1(p+n;m;\mathbf{\Omega}(\mathbf{I}+\mathbf{W}^{-\mathbf{1}})^{-\mathbf{1}})}{\Gamma_n(m)\Gamma_n(p+n-m)\mathrm{e}^{\mathsf{Tr}\{\mathbf{\Omega}\}}|\mathbf{W}|^{n-m}|\mathbf{I}+\mathbf{W}|^{p+n}} \tag{A.10}$$

and the distribution of its eigenvalues is given by

$$p(\mathbf{\Lambda}) = \frac{p!^n}{(m-n)!^n}\frac{\pi_n\mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Omega}\}}\mathcal{V}(\mathbf{\Lambda})|\mathbf{F}||\mathbf{\Lambda}|^{m-n}}{\Gamma_n(p+n-m)\mathcal{V}(\mathbf{\Omega})|\mathbf{I}+\mathbf{\Lambda}|^{1+p}} \tag{A.11}$$

where $\mathbf{\Omega} = \mathbf{M}^{\mathsf{H}}\mathbf{\Theta}^{-1}\mathbf{M}$, $(\mathbf{F})_{ij} = {}_1F_1(p+1; m-n+1; \lambda_j\omega_i/(1+\lambda_j))$, and $\omega_1, \ldots, \omega_n$ are the eigenvalues of $\mathbf{\Omega}$. This result has been obtained by applying (2.18) to [29, eq. (106)].

In the particular case where $\mathbf{\Omega}$ is a scalar matrix (i.e., $\mathbf{\Omega} = \omega\mathbf{I}$), matrix $\mathbf{W}$ is unitarily invariant since its pdf in (A.10) only depends on its eigenvalues $\mathbf{\Lambda}$. Indeed, $|\mathbf{W}| = |\mathbf{\Lambda}|$, $|\mathbf{I} + \mathbf{W}| = |\mathbf{I} + \mathbf{\Lambda}|$, and the generalized hypergeometric function ${}_1F_1(p+n; m; \omega(\mathbf{I} + \mathbf{W}^{-1})^{-1})$ only depends on the eigenvalues of its matrix argument, i.e., on $\mathbf{\Lambda}$. In such a case, the distribution of $\mathbf{\Lambda}$ can be obtained from (A.11) by applying the limit in (2.9) to the ratio $|\mathbf{F}|/\mathcal{V}(\mathbf{\Omega})$ and the property in (2.20). The result is reported in (2.37).

A.3 Proof of Theorem3

The proof of (2.38) and (2.40) follows from the application of [68, Theorem I] to (2.32) and (2.33), respectively.

The density given in (2.33) is an ordered eigenvalue distribution and the unordered eigenvalue distribution is obtained by dividing (2.33) by $n!$. Then, applying the Laplace determinant expansion, the unordered eigenvalues distribution becomes

$$\begin{aligned} p(\mathbf{\Lambda}) &= \frac{\pi_n^2 \Gamma_n(p+q)(1-\lambda_1)^{q-n}\lambda_1^{p-n-2}}{n!\Gamma_n(n)\Gamma_n(p)\Gamma_n(q)} \\ &\quad \cdot \sum_{i=1}^{n}\sum_{j=1}^{n}(-\lambda_1)^{i+j} \\ &\quad \cdot \prod_{k=2}^{n}(1-\lambda_k)^{q-n}\lambda_k^{q-n}|\bar{\mathbf{V}}(\mathbf{\Lambda})||\widetilde{\mathbf{V}}(\mathbf{\Lambda})| \end{aligned} \tag{A.12}$$

where $\bar{\mathbf{V}}(\mathbf{\Lambda})$ and $\widetilde{\mathbf{V}}(\mathbf{\Lambda})$ are $(n-1)\times(n-1)$ matrices obtained by deleting the first row and column from the Vandermonde matrix $\mathbf{V}(\mathbf{\Lambda})$ and its conjugate transpose, separately. The (i,j)-th entry of $\mathbf{V}(\mathbf{\Lambda})$ and its conjugate transpose are λ_i^{j-1} and λ_j^{i-1}, respectively. Thanks to [69, Corollary 1], the result in (2.40) can be obtained through integration over $n-1$ eigenvalues from λ_2 to λ_n. The final expressions are in both cases due to the definition of the scalar Beta function [32]. It should be noticed that the choice of λ_1 in (A.12) has no effect on the final result, since we started from an unordered eigenvalue distribution. Using the same approach, the proof of (2.38) is straightforward.

A.4 Proof of (3.2) and (3.4)

We first observe that for $m > n$ the matrix $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ does not have full rank and has $m-n$ zero eigenvalues. The n non-zero eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$, denoted by $\lambda_1,\ldots,\lambda_n$, are also the eigenvalues of $\mathbf{H}^{\mathsf{H}}\mathbf{H}$ and are the elements of the $n \times n$ diagonal matrix $\mathbf{\Lambda}$. We start by rewriting [35, eq. (38)] in the case $m > n$ and obtain

$$p(\mathbf{Y}) = \int \frac{\mathrm{e}^{-\|\mathbf{Y}\|^2} p(\mathbf{\Lambda})}{\pi^{mb}|\mathbf{I}+\gamma\mathbf{\Lambda}|^{\mathbf{b}}} \left[\int \mathrm{e}^{\mathsf{Tr}\{\widetilde{\mathbf{C}}\mathbf{U}^{\mathsf{H}}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\mathbf{U}\}} p(\mathbf{U}|\mathbf{\Lambda})\,\mathrm{d}\mathbf{U}\right]\mathrm{d}\mathbf{\Lambda} \tag{A.13}$$

where $\mathbf{U}$ is a unitary $m\times m$ matrix, $\widetilde{\mathbf{C}}$ is an $m\times m$ diagonal matrix whose elements are given by $(\widetilde{\mathbf{C}})_{jj} = c_j = \lambda_j\gamma/(1+\gamma\lambda_j)$, $j = 1,\ldots,m$, with $c_j = 0$, for $j = n+1,\ldots,m$. Since we assume that $\mathbf{W} = \mathbf{H}^{\mathsf{H}}\mathbf{H}$ is unitarily invariant, its eigenvalues do not depend on $\mathbf{U}$. Moreover, $\mathbf{U}$ is a Haar matrix. Then, $p(\mathbf{U}|\mathbf{\Lambda}) = \mathbf{p}(\mathbf{U})$. The inner integral over $\mathbf{U}$ can be solved using the Harish-Chandra-Itzykson-Zuber integral [70]

$$\int_{\mathcal{U}(m)} \mathrm{e}^{\mathsf{Tr}\{\widetilde{\mathbf{C}}\mathbf{U}^{\mathsf{H}}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\mathbf{U}\}} p(\mathbf{U})\,\mathrm{d}\mathbf{U} = \frac{\Gamma_m(m)|\mathbf{E}|}{\pi_m \mathcal{V}(\widetilde{\mathbf{C}})\mathcal{V}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}})}\,.$$

The elements of matrix $\mathbf{E}$ are given by $(\mathbf{E})_{ij} = \mathrm{e}^{y_i c_j}$, $i,j = 1,\ldots,m$ and y_i, $i = 1,\ldots,m$, are the eigenvalues of $\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$. Due to the fact that $c_j = 0$ for $j = n+1,\ldots,m$, we have $|\mathbf{E}| = 0$ and $\mathcal{V}(\widetilde{\mathbf{C}}) = 0$; thus the limit in (2.7) must be applied to the term $|\mathbf{E}|/\mathcal{V}(\widetilde{\mathbf{C}})$. We have

$$\lim_{c_{n+1},\ldots,c_m\to 0} \frac{|\mathbf{E}|}{\mathcal{V}(\widetilde{\mathbf{C}})} = \frac{\pi_m\Gamma_n(m)}{\pi_n\Gamma_m(m)} \frac{|\widetilde{\mathbf{E}}|}{\mathcal{V}(\mathbf{C})|\mathbf{C}|^{m-n}} \tag{A.14}$$

where $\widetilde{\mathbf{E}}$ is an $m\times m$ matrix whose elements are given by $(\widetilde{\mathbf{E}})_{ij} = \mathrm{e}^{y_i c_j}$ for $1 \le j \le n$, and $(\widetilde{\mathbf{E}})_{ij} = y_i^{j-n-1}$ for $n+1 \le j \le m$. Also, $\mathbf{C}$ is an $n \times n$ diagonal matrix whose elements are $(\mathbf{C})_{jj} = c_j = \lambda_j\gamma/(1+\gamma\lambda_j)$, $j = 1,\ldots,n$. Therefore, (A.13) can be rewritten as

$$p(\mathbf{Y}) = \frac{\Gamma_n(m)K(\mathbf{Y})}{\pi_n} \int \frac{p(\mathbf{\Lambda})|\widetilde{\mathbf{E}}|}{|\mathbf{I}+\gamma\mathbf{\Lambda}|^{\mathbf{b}}} \frac{|\mathbf{C}|^{n-m}}{\mathcal{V}(\mathbf{C})}\,\mathrm{d}\mathbf{\Lambda} \tag{A.15}$$

where $K(\mathbf{Y}) = \mathrm{e}^{-\|\mathbf{Y}\|^2}/(\mathcal{V}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}})\pi^{mb})$ was defined in (3.3). Since $c_j = \lambda_j\gamma/(1+\gamma\lambda_j)$, by applying the definition of the Vandermonde determinant, we get $\mathcal{V}(\mathbf{C}) = |\mathbf{I}+\gamma\mathbf{\Lambda}|^{\mathbf{1-n}}\mathcal{V}(\gamma\mathbf{\Lambda})$. Moreover, $|\mathbf{C}| = |\gamma\mathbf{\Lambda}||\mathbf{I}+\gamma\mathbf{\Lambda}|^{\mathbf{-1}}$. By substituting these results in (A.15), we obtain (3.4).

A.5 Proof of Proposition 1

We first observe that the matrix $\mathbf{H}$ in (3.7) can be written as $\mathbf{H} = \mathbf{H}_0/\sqrt{1+\kappa}$, where $\mathbf{H}_0 = \sqrt{\kappa}\bar{\mathbf{H}} + \widetilde{\mathbf{H}}$.

- For $m \leq n$ and $\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathsf{H}} = h\mathbf{I}$, the joint distribution of the ordered eigenvalues of $\mathbf{H}_0\mathbf{H}_0{}^{\mathsf{H}}$ is given by (2.28) where $\mu = \kappa h$, i.e.,

$$p_0(\mathbf{\Lambda}_0) = \frac{\pi_m^2 |\mathbf{\Lambda}_0|^n \mathcal{V}(\mathbf{\Lambda}_0) |\{\lambda_{0j}^{-i}{}_0F_1(\,;n-i+1;\kappa h\lambda_{0j})\}|}{\Gamma_m(m)\Gamma_m(n)\mathrm{e}^{\kappa hm+\mathsf{Tr}\{\mathbf{\Lambda}_0\}}}$$

where $(\lambda_{01},\ldots,\lambda_{0m}) = \mathrm{diag}(\mathbf{\Lambda_0})$. Then, the pdf of the ordered eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is given by

$$\begin{aligned} p(\mathbf{\Lambda}) &= (1+\kappa)^m p_0((1+\kappa)\mathbf{\Lambda}) \\ &= \frac{\pi_m^2 (1+\kappa)^{mn} |\mathbf{\Lambda}|^n |\mathbf{F}| \mathcal{V}(\mathbf{\Lambda})}{\Gamma_m(m)\Gamma_m(n)\mathrm{e}^{\kappa hm+(1+\kappa)\mathsf{Tr}\{\mathbf{\Lambda}\}}} \end{aligned} \tag{A.16}$$

where $(\mathbf{F})_{ij} = \lambda_j^{-i}{}_0F_1(\,;n-i+1;\kappa(1+\kappa)h\lambda_j)$, $i,j = 1,\ldots,m$. By substituting this equation in (3.2) and by applying the result in Appendix A.11, we obtain (3.8).

- For $m > n$, and for $\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}} = h\mathbf{I}$, we adopt a procedure similar to the one above. In this case, the pdf of the non-zero eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is given by (A.16) where n and m should be replaced by m and n, respectively. By substituting $p(\mathbf{\Lambda})$ in (3.4) and by applying the result in Appendix A.11, we obtain (3.9).

A.6 Proof of Proposition 2

For $m \leq n$, the distribution of the ordered eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is expressed as [42, eq. (9)]

$$p(\mathbf{\Lambda}) = \frac{\Gamma(\alpha-m+1)^m}{\Gamma(n-m+1)^m} \frac{\pi_m \mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Lambda}\}}}{\mathcal{V}((\mathbf{I}+\mathbf{\Omega})^{-1})} \frac{\mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}|\mathbf{F}|}{\Gamma_m(\alpha)|\mathbf{I}+\mathbf{\Omega}^{-1}|^\alpha} \tag{A.17}$$

with $(\mathbf{F})_{ij} = {}_1F_1(\alpha-m+1;n-m+1;\lambda_j/(1+\omega_i))$. When $\mathbf{\Omega} = \omega\mathbf{I}$, the expression of $p(\mathbf{\Lambda})$ can be derived from (A.17) by applying the limit in (2.9) and by using the property in (**??**). For simplicity, we define $\mathbf{\Theta} = (\mathbf{I}+\mathbf{\Omega})^{-\mathbf{1}} = \theta\mathbf{I}$ where $\theta = (1+\omega)^{-1}$. Then,

$$\begin{aligned} p(\mathbf{\Lambda}) &= \frac{\pi_m \Gamma(\alpha-m+1)^m \mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}}{\Gamma_m(\alpha)\Gamma(n-m+1)^m \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Lambda}\}}(1+1/\omega)^{m\alpha}} \\ &\quad \cdot \lim_{\mathbf{\Theta}\to\theta\mathbf{I}} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{\Theta})} \\ &= \frac{\pi_m^2 \mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^n |\widetilde{\mathbf{F}}|}{\Gamma_m(m)\Gamma_m(n)\mathrm{e}^{\mathsf{Tr}\{\mathbf{\Lambda}\}}(1+1/\omega)^{m\alpha}} \end{aligned} \tag{A.18}$$

where $(\widetilde{\mathbf{F}})_{ij} = \lambda_j^{-i}{}_1F_1(\alpha-i+1; n-i+1; \lambda_j/(1+\omega))$, $i,j = 1,\ldots,m$. The proposition statement follows by replacing (A.18) in (3.2). Similarly, when $m > n$ and $\mathbf{\Omega} = \omega\mathbf{I}$, the distribution of the eigenvalues of $\mathbf{H}^{\mathsf{H}}\mathbf{H}$ is given by

$$p(\mathbf{\Lambda}) = \frac{\pi_n^2 \mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Lambda}\}} \mathcal{V}(\mathbf{\Lambda}) |\mathbf{\Lambda}|^m}{\Gamma_n(m)(1+1/\omega)^{n\alpha}} \frac{|\widetilde{\mathbf{F}}|}{\Gamma_n(n)}$$

where $(\widetilde{\mathbf{F}})_{ij} = \lambda_j^{-i}{}_1F_1(\alpha - i + 1; m - i + 1; \lambda_j/(1+\omega))$, $i,j = 1,\ldots,n$. Again, the proposition statement is obtained by replacing the above equation in (3.4).

A.7 Proof of Proposition 4

We first observe that the matrix $\mathbf{H}$ in (3.13) can be written as $\mathbf{H} = \mathbf{H}_0/\sqrt{1+\kappa}$, where

$$\mathbf{H}_0 = \sqrt{\kappa}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2}\bar{\mathbf{H}}_s + \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2}\widetilde{\mathbf{H}}_s$$

- for $m \leq n$, and $\bar{\mathbf{H}}_s\bar{\mathbf{H}}_s{}^{\mathsf{H}} = h\mathbf{I}$, the distribution of the ordered eigenvalues of $\mathbf{H}_0\mathbf{H}_0{}^{\mathsf{H}}$ is given by (2.36)

$$\begin{aligned} p_0(\mathbf{\Lambda}_0) &= \frac{\pi_m^2 \mathrm{e}^{-h\kappa m/\theta}\Gamma_m(Ln+n)\mathcal{V}(\mathbf{\Lambda}_0)^2|\mathbf{\Lambda}_0|^{n-m}}{\Gamma_m(m)\Gamma_m(n)\Gamma_m(Ln)|\mathbf{I}+\mathbf{\Lambda}_0|^{Ln+n}} \\ &\quad \cdot{}_1F_1\left(Ln+n; n; \frac{h\kappa}{\theta}\mathbf{\Lambda}_0(\mathbf{I}+\mathbf{\Lambda}_0)^{-1}\right) \end{aligned}$$

where we set $\mu = h\kappa$. The distribution of the eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ can then be obtained as

$$\begin{aligned} p(\mathbf{\Lambda}) &= \tilde{\kappa}^m p_0(\tilde{\kappa}\mathbf{\Lambda}) \\ &= \frac{\pi_m^2 \tilde{\kappa}^{mn}\Gamma_m(Ln+n)|\mathbf{\Lambda}|^{n-m}\mathcal{V}^2(\mathbf{\Lambda})}{\Gamma_m(m)\Gamma_m(n)\Gamma_m(Ln)\mathrm{e}^{h\kappa m/\theta}|\mathbf{I}+\tilde{\kappa}\mathbf{\Lambda}|^{Ln+n}} \\ &\quad \cdot{}_1F_1\left(Ln+n; n; \frac{h\kappa\tilde{\kappa}}{\theta}\mathbf{\Lambda}(\mathbf{I}+\tilde{\kappa}\mathbf{\Lambda})^{-1}\right) \end{aligned} \tag{A.19}$$

where $\tilde{\kappa} = 1+\kappa$. By substituting the above expression in (3.2), and by exploiting the property [30, eq. (2.36)]

$${}_1F_1(a; b; \mathbf{\Psi}) = \frac{|\{{}_1F_1(a-m+j; b-m+j; \psi_i)\psi_i^{j-1}\}|}{\mathcal{V}(\mathbf{\Psi})},$$

which holds for any $m \times m$ Hermitian matrix $\mathbf{\Psi}$ with eigenvalues $\psi_1,\ldots,\psi_m$, we obtain (3.17).

- For $m > n$, the distribution of the eigenvalues of $\mathbf{H}_0\mathbf{H}_0^{\mathsf{H}}$ is given by (2.37):

$$p_0(\mathbf{\Lambda}_0) = \frac{\pi_n^2 \Gamma_n(Ln+n) \mathrm{e}^{-\omega n} |\mathbf{F}| |\mathbf{\Lambda}_0|^{m-n} \mathcal{V}(\mathbf{\Lambda}_0)}{\Gamma_n(n)\Gamma_n(m)\Gamma_n(Ln+n-m)|\mathbf{I}+\mathbf{\Lambda}_0|^{Ln+1}}$$

  where $\mathbf{\Omega} = \omega\mathbf{I} = \mathbf{M}^{\mathsf{H}}\mathbf{\Theta}^{-1}\mathbf{M}$. In our case we have $\mathbf{M} = \sqrt{\kappa}\bar{\mathbf{H}}_s$, thus $\omega\mathbf{I} = \kappa\bar{\mathbf{H}}_s{}^{\mathsf{H}}\mathbf{\Theta}^{-1}\bar{\mathbf{H}}_{\mathbf{s}}$. It follows that the matrix $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is unitarily invariant if $\bar{\mathbf{H}}_s{}^{\mathsf{H}}\mathbf{\Theta}^{-1}\bar{\mathbf{H}}_{\mathbf{s}} = \omega/\kappa\mathbf{I}$. The distribution of the eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ can then be obtained as

$$\begin{aligned} p(\mathbf{\Lambda}) &= \tilde{\kappa}^n p_0(\tilde{\kappa}\mathbf{\Lambda}) \\ &= \frac{\pi_n^2 \Gamma_n(Ln+n)|\mathbf{\Lambda}|^{m-n}\mathcal{V}(\mathbf{\Lambda})}{\Gamma_n(n)\Gamma_n(m)\Gamma_n(Ln+n-m)} \\ &\quad \cdot \frac{|\{\tilde{\lambda}_j^{n-i}{}_1F_1(Ln+n-i+1; m-i+1; \omega\tilde{\kappa}\tilde{\lambda}_j)\}|}{\tilde{\kappa}^{-nm}\mathrm{e}^{\omega n}|\mathbf{I}+\tilde{\kappa}\mathbf{\Lambda}|^{Ln+1}} \end{aligned}$$

  where $\tilde{\kappa} = 1+\kappa$ and $\tilde{\lambda}_j = \lambda_j/(1+\tilde{\kappa}\lambda_j)$. By substituting this expression in (3.4), we obtain (3.19).

A.8 Proof of Proposition 4

Given the above assumptions and considering that $\mathbf{X} = \sqrt{c}\mathbf{D}^{1/2}\mathbf{\Phi}$, $p(\mathbf{Y}|\mathbf{D})$ is given in as [2, eq. (53)]:

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\mathrm{e}^{-\|\mathbf{Y}\|^2} A}{\pi^{mb}|\mathbf{I}+\gamma c\mathbf{D}|^m} \tag{A.20}$$

where

$$\begin{aligned} A &= \int_{\mathcal{S}(b,n)} \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Delta}\mathbf{\Phi}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}\}} p(\mathbf{\Phi})\,\mathrm{d}\mathbf{\Phi} \\ &= \frac{1}{|\mathcal{S}(b,n)|} \int_{\mathcal{S}(b,n)} \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Delta}\mathbf{\Phi}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}\}}\,\mathrm{d}\mathbf{\Phi} \end{aligned}$$

and $\mathbf{\Delta} = \gamma c\mathbf{D}(\mathbf{I}+\gamma c\mathbf{D})^{-1}$. In [2, Appendix A], it is observed that the integral above is not an instance of the Harish-Chandra-Itzykson-Zuber (HCIZ) integral [70] since the $n \times b$ matrix $\mathbf{\Phi}$ is not a square matrix. In order to circumvent this problem, one has to extend matrix $\mathbf{\Phi}^{\mathsf{H}}$ to the unitary $b \times b$ Haar matrix $\widetilde{\mathbf{\Phi}}^{\mathsf{H}} = [\mathbf{\Phi}^{\mathsf{H}}, \mathbf{\Phi}_{\perp}{}^{\mathsf{H}}]$, where $\mathbf{\Phi}_{\perp}{}^{\mathsf{H}}$ is the orthogonal complement of $\mathbf{\Phi}^{\mathsf{H}}$ with respect to the unitary group $\mathcal{U}(b)$. Thus, following [2, Appendix A], we can write

$$A = \frac{1}{|\mathcal{S}(b,n)||\mathcal{U}(b-n)|} \int_{\mathcal{U}(b)} \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Delta}\mathbf{\Phi}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}\}}\,\mathrm{d}\widetilde{\mathbf{\Phi}}\,.$$

Next, the $n \times n$ diagonal matrix $\mathbf{\Delta} = \mathrm{diag}(\delta_{\mathbf{1}}, \ldots, \delta_{\mathbf{n}})$ can be extended to the $b \times b$ matrix $\widetilde{\mathbf{\Delta}} = \mathrm{diag}(\delta_1, \ldots, \delta_n, q_1, \ldots, q_{b-n})$ where the elements of $\mathbf{q} = [q_1, \ldots, q_{b-n}]$ are distinct and different from $\delta_1, \ldots, \delta_n$. The above integral can then be written as

$$A = \frac{1}{|\mathcal{S}(b,n)||\mathcal{U}(b-n)|} \lim_{\mathbf{q}\to 0} \int_{\mathcal{U}(b)} \mathrm{e}^{\mathsf{Tr}\{\widetilde{\mathbf{\Delta}}\widetilde{\mathbf{\Phi}}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\widetilde{\mathbf{\Phi}}^{\mathsf{H}}\}} \, \mathrm{d}\widetilde{\mathbf{\Phi}} \, . \tag{A.21}$$

We observe that the matrix $\mathbf{Y}^{\mathsf{H}}\mathbf{Y}$ has $b - m$ zero-eigenvalues, and its non-zero eigenvalues are the eigenvalues of $\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$. Since the HCIZ integral is a function of the eigenvalues of the matrices $\widetilde{\mathbf{\Delta}}$ and $\mathbf{Y}^{\mathsf{H}}\mathbf{Y}$, we replace the matrix $\mathbf{Y}^{\mathsf{H}}\mathbf{Y}$ with the $b \times b$ block diagonal matrix $\mathbf{\Psi} = \mathrm{diag}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}}, \mathbf{P})$ where $\mathbf{P}$ is diagonal and has diagonal entries $\mathbf{p} = [p_1, \ldots, p_{b-m}]$. Such elements are positive, distinct, and they are different from the eigenvalues of $\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$. In conclusion, we can write:

$$\begin{aligned} A &= \frac{1}{|\mathcal{S}(b,n)||\mathcal{U}(b-n)|} \lim_{\mathbf{q}\to\mathbf{0}} \lim_{\mathbf{p}\to\mathbf{0}} \int_{\mathcal{U}(b)} \mathrm{e}^{\mathsf{Tr}\{\widetilde{\mathbf{\Delta}}\widetilde{\mathbf{\Phi}}\mathbf{\Psi}\widetilde{\mathbf{\Phi}}^{\mathsf{H}}\}} \, \mathrm{d}\widetilde{\mathbf{\Phi}} \\ &= \frac{\Gamma_b(b)|\mathcal{U}(b)|}{\pi_b|\mathcal{S}(b,n)||\mathcal{U}(b-n)|} \lim_{\mathbf{q}\to 0} \lim_{\mathbf{p}\to\mathbf{0}} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{\Psi})\mathcal{V}(\widetilde{\mathbf{\Delta}})} \\ &= \frac{\Gamma_b(b)}{\pi_b} \lim_{\mathbf{q}\to 0} \lim_{\mathbf{p}\to\mathbf{0}} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{\Psi})\mathcal{V}(\widetilde{\mathbf{\Delta}})} \end{aligned} \tag{A.22}$$

where $(\mathbf{F})_{ij} = \mathrm{e}^{\psi_i \tilde{\delta}_j}$ and ψ_i and $\tilde{\delta}_j$ are the eigenvalues of $\mathbf{\Psi}$ and $\widetilde{\mathbf{\Delta}}$, respectively. In (A.22) we first used the HCIZ integral [70] and then the equality $|\mathcal{U}(b)| = |\mathcal{S}(b,n)||\mathcal{U}(b-n)|$. Then, we apply twice the limit in (2.9) and obtain:

$$\begin{aligned} \lim_{\mathbf{q}\to 0} \lim_{\mathbf{p}\to\mathbf{0}} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{\Psi})\mathcal{V}(\widetilde{\mathbf{\Delta}})} &= \frac{\pi_b \Gamma_m(b)|\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{m-b}}{\pi_m \Gamma_b(b)\mathcal{V}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}})} \\ &\quad \cdot \frac{\pi_b \Gamma_n(b)|\widehat{\mathbf{F}}||\mathbf{\Delta}|^{n-b}}{\pi_n \Gamma_b(b)\mathcal{V}(\mathbf{\Delta})} \end{aligned} \tag{A.23}$$

where for $m \leq n$,

$$(\widehat{\mathbf{F}})_{ij} = \begin{cases} \mathrm{e}^{y_i \delta_j} & i = 1, \ldots, m; j = 1, \ldots, n \\ y_i^{b-j} & i = 1, \ldots, m; j = n+1, \ldots, b \\ \delta_j^{b-i} & i = m+1, \ldots, b, j = 1, \ldots, n \\ (b-i)! & i = j; j = n+1, \ldots, b \\ 0 & \text{elsewhere} \, , \end{cases} \tag{A.24}$$

while for $m > n$,

$$(\widehat{\mathbf{F}})_{ij} = \begin{cases} \mathrm{e}^{y_i \delta_j} & i = 1, \ldots, m; j = 1, \ldots, n \\ y_i^{b-j} & i = 1, \ldots, m; j = n+1, \ldots, b \\ \delta_j^{b-i} & i = m+1, \ldots, b, j = 1, \ldots, n \\ (b-i)! & i = j; j = m+1, \ldots, b \\ 0 & \text{elsewhere.} \end{cases} \tag{A.25}$$

In summary,

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\pi_b \Gamma_m(b) \Gamma_n(b) K(\mathbf{Y}) |\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{m-b} |\widehat{\mathbf{F}}| |\boldsymbol{\Delta}|^{n-b}}{\pi_m \pi_n \Gamma_b(b) \mathcal{V}(\boldsymbol{\Delta}) |\mathbf{I} + \gamma c \mathbf{D}|^m} \tag{A.26}$$

where $K(\mathbf{Y})$ was defined in (3.3).

We now focus on the case $m > n$ and compute the determinant $|\widehat{\mathbf{F}}|$. Note that $\widehat{\mathbf{F}}$ can be written as

$$\widehat{\mathbf{F}} = \begin{bmatrix} \widehat{\mathbf{F}}_1 & \widehat{\mathbf{F}}_2 \\ \widehat{\mathbf{F}}_3 & \widehat{\mathbf{F}}_4 \end{bmatrix}$$

where $\widehat{\mathbf{F}}_1$ is of size $m \times m$, $\widehat{\mathbf{F}}_2$ $m \times (b-m)$, $\widehat{\mathbf{F}}_3$ $(b-m) \times m$, and $\widehat{\mathbf{F}}_4$ $(b-m) \times (b-m)$. By using the property of the determinant of block matrices [71], we have:

$$|\widehat{\mathbf{F}}| = |\widehat{\mathbf{F}}_4| |\widehat{\mathbf{T}}| \,,$$

where $\widehat{\mathbf{T}} = \widehat{\mathbf{F}}_1 - \widehat{\mathbf{F}}_2 \widehat{\mathbf{F}}_4^{-1} \widehat{\mathbf{F}}_3$. In our case, $\widehat{\mathbf{F}}_4$ is diagonal (see the definition of $\widehat{\mathbf{F}}$ in (A.25)) and $|\widehat{\mathbf{F}}_4| = \prod_{i=0}^{b-m-1} i!$. Moreover, we have $(\widehat{\mathbf{F}}_2 \widehat{\mathbf{F}}_4^{-1} \widehat{\mathbf{F}}_3)_{ij} = \sum_{k=0}^{b-m-1} (y_i \delta_j)^k / k!$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$, and $(\widehat{\mathbf{F}}_2 \widehat{\mathbf{F}}_4^{-1} \widehat{\mathbf{F}}_3)_{ij} = 0$ otherwise. It follows that for $i = 1, \ldots, m$

$$(\widehat{\mathbf{T}})_{ij} = \begin{cases} \mathrm{e}^{y_i \delta_j} - \sum_{k=0}^{b-m-1} \frac{(y_i \delta_j)^k}{k!} & j = 1, \ldots, n \\ y_i^{b-j} & j = n+1, \ldots, m \,. \end{cases}$$

Note that, for $i = 1,\ldots,m$ and $j = 1,\ldots,n$,

$$\begin{aligned}
(\widehat{\mathbf{T}})_{ij} &= \mathrm{e}^{y_i\delta_j} - \sum_{k=0}^{b-m-1} \frac{(y_i\delta_j)^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{(y_i\delta_j)^k}{k!} - \sum_{k=0}^{b-m-1} \frac{(y_i\delta_j)^k}{k!} \\
&= \sum_{k=b-m}^{\infty} \frac{(y_i\delta_j)^k}{k!} \\
&= \sum_{h=0}^{\infty} \frac{(y_i\delta_j)^{b-m+h}}{(b-m+h)!} \\
&= (y_i\delta_j)^{b-m} \sum_{h=0}^{\infty} \frac{(y_i\delta_j)^h h!}{(b-m+h)!h!} \\
&= \frac{(y_i\delta_j)^{b-m}}{(b-m)!} \sum_{h=0}^{\infty} \frac{(1)_h (y_i\delta_j)^h}{(b-m+1)_h h!} \\
&= \frac{(y_i\delta_j)^{b-m}}{(b-m)!} {}_1F_1(1; b-m+1; y_i\delta_j)
\end{aligned}$$

since $h! = (1)_h$ and $(b-m+h)! = (b-m+1)_h(b-m)!$. Also, for $i = 1,\ldots,m$ and $j = n+1,\ldots,m$, $(\widehat{\mathbf{T}})_{ij} = y_i^{b-j} = y_i^{b-m} y_i^{m-j}$.

As a consequence, the matrix $\widehat{\mathbf{T}}$ can be rewritten as $\widehat{\mathbf{T}} = \mathbf{LGR}$, where $\mathbf{L}$ and $\mathbf{R}$ are diagonal $m \times m$ matrices given by, respectively, $\mathbf{L} = \mathrm{diag}(y_1^{b-m},\ldots,y_m^{b-m})$, and

$$\mathbf{R} = \mathrm{diag}(\delta_1^{b-m}/(b-m)!,\ldots,\delta_n^{b-m}/(b-m)!,1,\ldots,1)\,.$$

Furthermore, $\mathbf{G}$ is an $m \times m$ matrix whose elements, for i=1,…,m, are given by

$$(\mathbf{G})_{ij} = \begin{cases} {}_1F_1(1; b-m+1; y_i\delta_j) & j = 1,\ldots,n \\ y_i^{m-j} & j = n+1,\ldots,m\,. \end{cases}$$

Thus, we have:

$$\begin{aligned}
|\widehat{\mathbf{F}}| &= |\widehat{\mathbf{F}}_4||\widehat{\mathbf{T}}| \\
&= |\mathbf{L}||\mathbf{G}||\mathbf{R}| \prod_{i=0}^{b-m-1} i! \\
&= \frac{|\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{b-m}|\mathbf{G}||\boldsymbol{\Delta}|^{b-m}}{(b-m)!^n} \prod_{i=0}^{b-m-1} i!\,. \qquad \text{(A.27)}
\end{aligned}$$

In conclusion, by substituting (A.27) in (A.26), we get (3.25).

For $m \leq n$, a similar procedure can be used to compute the determinant $|\widehat{\mathbf{F}}|$. In this case,

$$|\widehat{\mathbf{F}}| = \frac{|\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{b-n}|\mathbf{G}||\boldsymbol{\Delta}|^{b-n}}{(b-n)!^m} \prod_{i=0}^{b-n-1} i!\,.$$

Again, by substituting the above expression in (A.26), we get (3.24). Here, however, the expression of $(\mathbf{G})_{ij}$ changes as follows:

$$(\mathbf{G})_{ij} = \begin{cases} {}_1F_1(1; b-n+1; y_i\delta_j) & i = 1,\ldots,m \\ \delta_j^{n-i} & i = m+1,\ldots,n \end{cases}$$

and for $j = 1,\ldots,n$.

A.9 Proof of Proposition 6

The law of the output of a channel, as the one in (3.1), conditioned on the input power allocation $\mathbf{D}$, is reported in [2, eq. (58)], i.e.,

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\Gamma_n(b)K(\mathbf{Y})|\mathbf{I} + c\gamma\mathbf{D}|^{b-m-1}|\mathbf{F}|}{\pi_n\gamma_n|\gamma c\mathbf{D}|^{b-n}\mathcal{V}(c\mathbf{D})}\,, \tag{A.28}$$

where $K(\mathbf{Y}) = \mathrm{e}^{-\|\mathbf{Y}\|^2}/(\pi^{mb}\mathcal{V}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}}))$ and $(\mathbf{F})_{ij} = \exp\left(\frac{c\gamma y_i d_j}{1+c\gamma d_j}\right)$, $i = 1,\ldots,n$,$j = 1,\ldots,b$, and $\mathbf{F}_{ij} = y_i^{b-j}$, $i = n+1,\ldots,b$,$j = 1,\ldots,b$.

In order to take average of (A.28), we first write $|\mathbf{F}|$ as the product of two determinants. Indeed, we partition $\mathbf{F}$ as

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 \\ \mathbf{F}_3 & \mathbf{F}_4 \end{bmatrix} \tag{A.29}$$

where $(\mathbf{F}_4)_{ij} = y_{n+i}^{b-n-j}$ i,$j = 1,\ldots,b-n$, and $\mathbf{F}_1$ is the principle $n \times n$ submatrix of $\mathbf{F}$. Applying the property of the determinant of block matrices [71] to (A.29), we obtain

$$|\mathbf{F}| = |\mathbf{F}_4||\mathbf{T}|\,, \tag{A.30}$$

where $\mathbf{T} = \mathbf{F}_1 - \mathbf{F}_2\mathbf{F}_4^{-1}\mathbf{F}_3$. We notice that $|\mathbf{F}_4|$ is independent of $\mathbf{D}$, and the matrix

$\mathbf{T}$ has the same size as $\mathbf{D}$. For $m > n$, $p(\mathbf{D})$ is given by (2.33). We then get

$$\begin{aligned} p(\mathbf{Y}) &= \int p(\mathbf{Y}|\mathbf{D})p(\mathbf{D})\,\mathrm{d}\mathbf{D} \\ &= \frac{\Gamma_n(b)K(\mathbf{Y})|\mathbf{F}_4|}{\pi_n\gamma_n}\int \frac{|\mathbf{I}+c\gamma\mathbf{D}|^{b-m-1}|\mathbf{T}|p(\mathbf{D})}{|\gamma c\mathbf{D}|^{b-n}\mathcal{V}(c\mathbf{D})}\,\mathrm{d}\mathbf{D} \\ &= \frac{\pi_n\Gamma_n(b)\Gamma_n(m)(\gamma c)^{n(n-b)}K(\mathbf{Y})|\mathbf{F}_4|}{\gamma_n\Gamma_n(n)c^{n(n-1)/2}\Gamma_n(b-n)\Gamma_n(n+m-b)} \\ &\quad \cdot\int \frac{|\mathbf{I}+c\gamma\mathbf{D}|^{b-m-1}|\mathbf{T}|}{|\mathbf{I}-\mathbf{D}|^{b-m}|\mathbf{D}|^{n}}\mathcal{V}(\mathbf{D})\,\mathrm{d}\mathbf{D}\,, \\ &= \frac{\pi_n\Gamma_n(b)\Gamma_n(m)(\gamma c)^{n(n-b)}K(\mathbf{Y})|\mathbf{F}_4||\mathbf{Z}|}{\gamma_n c^{n(n-1)/2}\Gamma_n(n)\Gamma_n(b-n)\Gamma_n(n+m-b)} \end{aligned} \tag{A.31}$$

where

$$\begin{aligned} (\mathbf{Z})_{ij} &= \int_0^1 \frac{(1+c\gamma x)^{b-m-1}x^{i-1-n}}{(1-x)^{b-m}} \\ &\quad \cdot\left[\exp\left(\frac{c\gamma y_i x}{1+c\gamma x}\right) - (\mathbf{F}_2\mathbf{F}_4^{-1}\mathbf{F}_3)_{ij}\right] \end{aligned}$$

has been obtained by using the result in Appendix A.11, and

$$(\mathbf{F}_2\mathbf{F}_4^{-1}\mathbf{F}_3)_{ij} = \sum_{\ell,k=1}^{b-n}(\mathbf{F}_4^{-1})_{\ell k}\exp\left(\frac{c\gamma x y_{\ell+n}}{1+c\gamma x}\right)y_j^{b-k-n}\,.$$

A.10 Proof of Proposition 7

Conditioned on the input $\mathbf{X}$, the output $\mathbf{Y}$ is complex Gaussian and has i.i.d. rows, so that the evaluation of the differential entropy can be carried out by considering just an arbitrary row, $\mathbf{y}$, of $\mathbf{Y}$ and, then, scaling the result by m. We note that $\mathbf{y}$ is multivariate Gaussian distributed with covariance equal to $(\mathbf{I}+\gamma\mathbf{X}^{\mathsf{H}}\mathbf{X})$. Thus, considering the optimal input matrix, $\mathbf{X} = \sqrt{c}\mathbf{D}^{1/2}\mathbf{\Phi}$ and conditioning on it, the differential entropy is given by:

$$h(\mathbf{y}|\mathbf{X}) = b\log_2(\pi e) + n\mathbb{E}\left[\log_2\left(1+c\gamma\delta\right)\right]\,, \tag{A.32}$$

with δ being distributed as a single unordered eigenvalue of the matrix $\mathbf{D}$. By using (2.33) and considering $p = b-n$ and $q = m+n-b$, $p(\mathbf{D})$ reads as

$$p(\mathbf{D}) = \frac{\pi_n^2\Gamma_n(m)|\mathbf{I}-\mathbf{D}|^{m-b}|\mathbf{D}|^{b-2n}\mathcal{V}^2(\mathbf{D})}{\Gamma_n(n)\Gamma_n(b-n)\Gamma_n(m+n-b)}\,. \tag{A.33}$$

By exploiting the result given in Proposition 3 and by denoting the constant terms in the above expression by K, we obtain:

$$p(\delta) = \frac{K}{n} \sum_{i,j=1}^{n} \delta^{(b-2n+i+j-2)} (1-\delta)^{(m-b)} a_{ij} \tag{A.34}$$

with a_{ij} being defined as in the above proposition. The integral in (A.32) can be solved by resorting to partial integration. Indeed, taking $\log_2(1+c\gamma\delta)$ as the primitive factor and recalling that $(1-\delta)^{n-m} = \sum_{\ell=0}^{n-m} \binom{n-m}{\ell} (-1)^\ell \delta^\ell$, by virtue of [72, 3.194.1], we obtain

$$\mathbb{E}\left[\log_2(1+\gamma\delta)\right] = \frac{K}{n} \sum_{i,j=1}^{n} a_{ij} \sum_{\ell=0}^{m-b} \binom{m-b}{\ell} \frac{(-1)^\ell}{s_{i,j,\ell}-1}$$
$$\cdot \left[\log_2(1+c\gamma) - c\gamma \frac{{}_2F_1\left(1, s_{i,j,\ell}; s_{i,j,\ell}+1; -\gamma\right)}{\ln(2)(s_{i,j,\ell})} \right]$$

where $s_{i,j,\ell} = b-2n+i+j+\ell$. Then, using this expression in (A.32), we get (3.30).

A.11 Lemma 2 in [1]

Consider a function $\xi(x)$, an arbitrary $n \times n$ matrix $\mathbf{\Phi}(\mathbf{x})$ such that $(\mathbf{\Phi})_{\mathbf{ij}} = \phi_{\mathbf{i}}(\mathbf{x_j})$, and an arbitrary $m \times m$ matrix $\mathbf{\Psi}$, $m \geq n$, whose elements are given by

$$(\mathbf{\Psi})_{\mathbf{ij}} = \begin{cases} \psi_i(x_j) & 1 \leq i \leq m, 1 \leq j \leq n \\ c_{ij} & 1 \leq i \leq m, n+1 \leq j \leq m \end{cases}$$

where c_{ij} are constant. Then, the following identity holds:

$$\int_{[a,b]^n} |\mathbf{\Phi}(\mathbf{x})||\mathbf{\Psi}(\mathbf{x})| \prod_{\mathbf{k=1}}^{\mathbf{n}} \xi(\mathbf{x_k})\, \mathrm{d}\mathbf{x} = \mathbf{n}!|\mathbf{\Xi}| \tag{A.35}$$

where, for $1 \leq i \leq m$,

$$(\mathbf{\Xi})_{\mathbf{ij}} = \begin{cases} \int_a^b \psi_i(x)\phi_j(x)\xi(x)\,\mathrm{d}x & 1 \leq j \leq n \\ c_{ij} & n+1 \leq j \leq m\,. \end{cases}$$

For the specific case $m = n$, this result appears in [69, Corollary II].

Appendix B

Publications

The following is a list of papers in refereed journals and conference proceedings that have been published or submitted during my Ph.D. period.

1. S. Zhou, G. Alfano, A. Nordio, and C. F. Chiasserini, "Ergodic capacity analysis of MIMO relay network over Rayleigh-Rician channels," *IEEE Communications Letters*, to appear, 2015.

2. G. Alfano, C. F. Chiasserini, A. Nordio and S. Zhou, "Closed-form output statistics of MIMO block-fading channels," *IEEE Transaction on Information Theory*, Vol. 60, No. 12, pp. 7782-7797, Nov. 2014.

3. S. Zhou, A. Nordio and C. F. Chiasserini, "Estimation quality of high-dimensional fields in wireless sensor networks," *The Sixth International Conference on Advances in Satellite and Space Communications (SPACOMM 2014)*, Nice, France, 2014.

4. F. Malandrino, C. Casetti, C. F. Chiasserini, S. Zhou, "Real-time scheduling for content broadcasting in LTE," *IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS 2014)*, Paris, France, 2014.

5. S. Zhou, G. Alfano, C. F. Chiasserini and A. Nordio, "Characterization of output signals for MIMO block-fading channels with imperfect CSI," *IEEE International Symposium on Wireless Communication Systems (ISWCS 2013)*, Ilmenau, Germany, 2013.

6. G. Alfano, C. F. Chiasserini, A. Nordio and S. Zhou, "Output statistics of MIMO channels with general input distribution," *IEEE International Symposium on Information Theory Proceedings (ISIT 2013)*, Istanbul, Turkey, 2013.

7. G. Alfano, C. F. Chiasserini, A. Nordio and S. Zhou, "Information densities for block-fading MIMO channels," *IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC 2013)*, Turin, Italy, 2013.

Bibliography

[1] H. Shin, M. Win, J. Lee, M. Chiani, "On the capacity of doubly correlated MIMO channels," in *IEEE Trans. Wireless Commun.*, v. 5, n. 8, pp. 2253–2265, Aug. 2006.

[2] W. Yang, G. Durisi, E. Riegler, "On the capacity of large-MIMO block-fading channels," in *Selected Areas in Communications, IEEE Journal on*, v. 31, n. 2, pp. 117–132, Apr. 2013.

[3] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," in *Europ. Trans. Commun.*, v. 10, n. 6, pp. 585–595, Nov.-Dec. 1999.

[4] G. J. Foschini, M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," in *Wireless Pers. Commun.*, v. 6, n. 3, pp. 311–335, 1998.

[5] H. Bolcskei, D. Gesbert, A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," in *IEEE Trans. Commun.*, v. 50, n. 2, pp. 225–234, Feb. 2002.

[6] L. Zheng, D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," in *IEEE Trans. Inform. Theory*, v. 49, n. 5, pp. 1073–1096, May 2003.

[7] A. Goldsmith, *Wireless Communications,* Cambridge university press, 2005.

[8] T. L. Marzetta, B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," in *IEEE Trans. Inform. Theory*, v. 45, n. 1, pp. 139–157, 1999.

[9] N. Jindal, A. Lozano, "A unified treatment of optimum pilot overhead in multipath fading channels," in *IEEE Trans. Commun.*, v. 58, n. 10, pp. 2939–2948, Oct. 2010.

[10] S. K. Jayaweera, H. V. Poor, "Capacity of multiple-antenna systems with both receiver and transmitter channel state information," in *IEEE Trans. Inform. Theory*, v. 49, n. 10, pp. 2697–2709, Oct. 2003.

[11] G. J. Foschini, "Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas," in *Bell Labs Tech. J.*, v. 1, n. 2, pp. 41–59, 1996.

[12] D. P. Palomar, J. M. Cioffi, M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," in *IEEE Trans. Signal Processing*, v. 51, n. 9, pp. 2381–2401, Sept. 2003.

[13] S. K. Jayaweera, H. V. Poor, "On the capacity of multi-antenna systems in the presence of Rician fading," in *IEEE Trans. Wireless Commun.*, v. 4, n. 3, pp. 1102–1111, May 2005.

[14] M. kang, M. S. Alouini, "Impact of correlation on the capacity of MIMO channels," in *IEEE Trans. Wireless Commun.*, v. 5, n. 1, pp. 112–122, Jan. 2006.

[15] T. Yoo, A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," in *IEEE Trans. Inform. Theory*, v. 52, n. 2, pp. 2203–2214, May 2006.

[16] H. Shin, J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," in *IEEE Trans. Inform. Theory*, v. 49, n. 10, pp. 2636–2647, Oct. 2003.

[17] M. Chiani, M. Z. Win, A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," in *IEEE Trans. Inform. Theory*, v. 49, n. 10, pp. 2363–2371, Oct. 2003.

[18] G. Alfano, A. M. Tulino, S. Verdú, "Capacity of MIMO channels with one-sided correlation," in *IEEE Int. Symposium on Spread Spectrum Techniques and Applications (ISSSTA)*, Sydney, Australia, August 2004, pp. 515–519.

[19] G. Alfano, A. Lozano, A. M. Tulino, S. Verdú, "Mutual Information and Eigenvalue Distribution of MIMO Ricean Channels," in *IEEE Int. Symposium on Information Theory and its Applications (ISITA)*, Parma, Italy, Oct. 2004.

[20] S. Jin, X. Gao, X. You, "On the Ergodic Capacity of Rank-1 Ricean-Fading MIMO Channels," in *IEEE Trans. Inform. Theory*, v. 53, n. 2, pp. 502–517, Feb. 2007.

[21] Q. Zhang, X. Cui, X. Li, "Very tight capacity bounds for MIMO-correlated Rayleigh-fading channels," in *IEEE Trans. Wireless Commun.*, v. 4, n. 2, pp. 681–688, Mar. 2005.

[22] M. R. McKay, I. B. Collings, "General capacity bounds for spatially correlated rician MIMO channels," in *IEEE Trans. Inform. Theory*, v. 51, n. 9, pp. 3121–3145, Sept. 2005.

[23] G. Taricco, E. Riegler, "On the ergodic capacity of correlated rician fading MIMO channels with interference," in *IEEE Trans. Inform. Theory*, v. 49, n. 10, pp. 2636–2647, Oct. 2003.

[24] V. Tarokh, N. Seshadri, A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," in *IEEE Trans. Inform. Theory*, v. 44, n. 2, pp. 744–765, Mar. 1998.

[25] ——, "Space-time block codes from orthogonal designs," in *IEEE Trans. Inform. Theory*, v. 45, n. 5, pp. 1456–1467, July 1999.

[26] S. Alamouti, "A simple transmit diversity technique for wireless communications," in *Selected Areas in Communications, IEEE Journal on*, v. 16, n. 8, pp. 1451–1458, Oct. 1998.

[27] T. K. Lo, "Maximum ratio transmission," in *IEEE Trans. Commun.*, v. 47, n. 10, pp. 1458–1461, Oct. 1999.

[28] P. A. Dighe, R. K. Mallik, S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," in *IEEE Trans. Commun.*, v. 51, n. 4, pp. 694–703, Apr. 2003.

[29] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," in *The Annals of Mathematical Statistics*, v. 35, n. 2, pp. 475–501, June 1964.

[30] M. R. McKay, "Random matrix theory analysis of multiple antenna communication systems," Tesi di dottorato, University of Sydney, Oct. 2006.

[31] A. Ghaderipoor, C. Tellambura, A. Paulraj, "On the application of character expansions for MIMO capacity analysis," in *IEEE Trans. Inform. Theory*, v. 58, n. 5, pp. 2950–2962, May 2012.

[32] M. Abramowitz, I. A. Stegun, *et al.*, *Handbook of mathematical functions,* Dover New York, 1972.

[33] A. Constantine, "Some non-central distribution problems in multivariate analysis," in *The Annals of Mathematical Statistics*, v. 34, n. 4, pp. 1270–1285, Dec. 1963.

[34] A. M. Tulino, S. Verdú, *Random matrix theory and wireless communications,* Now Publishers Inc., 2004.

[35] F. Rusek, A. Lozano, N. Jindal, "Mutual information of IID complex Gaussian signals on block Rayleigh-faded channels," in *IEEE Trans. Inform. Theory*, v. 58, n. 1, pp. 331–340, Jan. 2012.

[36] A. Gupta, D. Nagar, A. Vélez-Carvajal, "Unitary invariant and residual independent matrix distributions," in *Computational and Applied Mathematics*, v. 28, n. 1, pp. 63–86, 2009.

[37] Y. Polyanskiy, H. V. Poor, S. Verdú, "Channel coding rate in the finite blocklength regime," in *IEEE Trans. Inform. Theory*, v. 56, n. 5, pp. 2307–2359, May 2010.

[38] B. Hassibi, T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: the received signal density in closed form," in *IEEE Trans. Inform. Theory*, v. 48, n. 6, pp. 1473–1484, June 2002.

[39] L. Zheng, D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," in *IEEE Trans. Inform. Theory*, v. 48, n. 2, pp. 359–383, Feb. 2002.

[40] S. Moser, "The fading number of IID MIMO Gaussian fading channels with a scalar line-of-sight component," in *Proceedings Forty-Fifth Allerton Conference on Communication, Control and Computing*, Monticello, IL, USA, 2007.

[41] F. Bohagen, P. Orten, G. E. Oien, "Design of optimal high-rank line-of-sight MIMO channels," in *IEEE Trans. Wireless Commun.*, v. 6, n. 4, pp. 1420–1425, 2007.

[42] G. Alfano, A. D. Maio, A. M. Tulino, "A theoretical framework for LMS MIMO communication systems performance analysis," in *IEEE Trans. Inform. Theory*, v. 56, n. 11, pp. 5614–5630, 2010.

[43] J. S. Kwak, J. Andrews, A. Lozano, "MIMO capacity in correlated interference-limited channels," in *IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.

[44] M. Gursoy, H. V. Poor, S. Verdú, "Spectral efficiency of peak power limited Rician block-fading channels," in *IEEE International Symposium on Information Theory (ISIT)*, Chicago, IL, USA, June 2004.

[45] G. Taricco, G. Coluccia, "Optimum receiver design for correlated rician fading MIMO channels with pilot-aided detection," in *Selected Areas in Communications, IEEE Journal on*, v. 25, n. 7, pp. 1311–1321, Sep. 2007.

[46] K. Liolis, J. Gómez-Vilardebó, E. Casini, A. I. Pérez-Neira, "Statistical modeling of dual-polarized MIMO land mobile satellite channels," in *IEEE Trans. Commun.*, v. 58, n. 11, pp. 3077–3083, Nov. 2010.

[47] D. Hösli, A. Lapidoth, "The capacity of a MIMO Ricean channel is monotonic in the singular values of the mean," in *Proceedings Fifth ITG Conf. Source and Channel Coding*, Jan. 2004.

[48] A. K. Nagar, L. Cardeno, "Matrix-variate Kummer-Gamma distribution," in *Random Operators and Stochastic Equations*, v. 9, n. 3, pp. 208–217, 2001.

[49] A. Lapidoth, S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," in *IEEE Trans. Inform. Theory*, v. 49, n. 10, pp. 2426–2467, Oct. 2003.

[50] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," in *IEEE Trans. Wireless Commun.*, v. 9, n. 11, pp. 3590–3600, 2010.

[51] T. Yoo, A. J. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," in *IEEE Trans. Inform. Theory*, v. 52, n. 5, pp. 2203–2214, May 2006.

[52] G. Taricco, "Asymptotic mutual information statistics of separately correlated rician fading MIMO channels," in *IEEE Trans. Inform. Theory*, v. 54, n. 8, pp. 3490–3504, Aug. 2008.

[53] A. Aubry, I. Esnaola, A. M. Tulino, S. Venkatesan, "Achievable rate region for Gaussian MIMO MAC with partial CSI," in *IEEE Trans. Inform. Theory*, v. 59, n. 7, pp. 4139–4170, July 2013.

[54] M. Ding, S. D. Blostein, W. H. Mow, C. Siriteanu, "A general framework for MIMO transceiver design with imperfect CSI and transmit correlation," in *IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Tokyo, Japan, September 2009.

[55] J. Wagner, B. Rankov, A. Wittneben, "On the asymptotic capacity of the Rayleigh fading amplify-and-forward MIMO relay channel," in *IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.

[56] V. I. Morgenshtern, H. Bölcskei, "Crystallization in large wireless networks," in *IEEE Trans. Inform. Theory*, v. 53, n. 10, pp. 3319–3349, Oct. 2007.

[57] S. Jin, M. McKay, C. Zhong, K.-K. Wong, "Ergodic capacity analysis of amplify and forward MIMO dual-hop systems," in *IEEE Trans. Inform. Theory*, v. 56, n. 5, pp. 2204–2224, May 2010.

[58] L. Jayasinghe, N. Rajatheva, P. Dharmawansa, M. L. Aho, "Noncoherent amplify-and-forward MIMO relaying with OSTBC over Rayleigh-Rician fading channels," in *IEEE Trans. Veh. Technol.*, v. 62, n. 4, pp. 1610–1622, May 2013.

[59] M. Lin, K. An, J. Ouyang, Y. Huang, M. Li, "Effect of beamforming on multi-antenna two hop asymmetric fading channels with fixed gain relays," in *Progress In Electromagnetics Research*, v. 133, pp. 367–390, 2013.

[60] M. Kang, M.-S. Alouini, "Capacity of MIMO rician channels," in *IEEE Trans. Wireless Commun.*, v. 5, n. 1, pp. 112–122, Jan. 2006.

[61] P. A. Dighe, R. K. Mallik, S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," in *IEEE Trans. Commun.*, v. 51, n. 4, pp. 694–703, Apr. 2003.

[62] P. Jayasinghe, L. K. S. Jayasinghe, M. Juntti, M. Latva-Aho, "Performance analysis of optimal beamforming in fixed-gain AF MIMO relaying over asymmetric fading channels," in *IEEE Trans. Commun.*, v. 62, n. 4, pp. 1201–1217, Apr. 2014.

[63] G. Amarasuriya, C. Tellambura, M. Ardakani, "Performance analysis of hop-by-hop beamforming for dual-hop MIMO AF relay networks," in *IEEE Trans. Commun.*, v. 60, n. 7, pp. 1823–1837, July 2012.

[64] N. Yang, M. Elkashlan, J. Yuan, "Impact of opportunistic scheduling on cooperative dual-hop relay networks," in *IEEE Trans. Commun.*, v. 59, n. 3, pp. 689–694, Mar. 2011.

[65] Y. Huang, F. Al-Qahtani, C. Zhong, Q. Wu, J. Wang, H. Alnuweiri, "Performance analysis of multiuser multiple antenna relaying networks with co-channel interfernce and feedback delay," in *IEEE Trans. Commun.*, v. 62, n. 1, pp. 59–73, Jan. 2014.

[66] C. G. Khatri, "On certain distribution problems based on positive definite quadratic functions in normal vectors," in *The Annals of Mathematical Statistics*, v. 37, n. 2, pp. 467–479, 1966.

[67] M. R. McKay, I. Collings, "General capacity bounds for spatially correlated Rician MIMO channels," in *IEEE Trans. Inform. Theory*, v. 51, n. 9, pp. 3121–3145, Sep. 2005.

[68] G. Alfano, A. Tulino, A. Lozano, S. Verdú, "Eigenvalue statistics of finite-dimensional random matrices for MIMO wireless communications," in *IEEE ICC*, Istanbul, Turkey, June 2006.

[69] M. Chiani, M. Z. Win, A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," in *IEEE Trans. Inform. Theory*, v. 49, n. 10, pp. 2363–2371, Oct. 2003.

[70] C. Itzykson, J. B. Zuber, "The planar approximation. II," in *J. of Mathematical Physics*, v. 21, n. 3, pp. 411–421, 1980.

[71] R. A. Horn, C. Johnson, *Matrix Analysis,* Cambridge University Press, 1985.

[72] I. S. Gradshteyn, I. M. Ryzhik, *Table of Integrals, Series, and Products,* Academic Press, New York, 1980.